Charles University in Prague

Faculty of Science

Study programme: Developmental and Cell Biology

Mgr. Monika Šťastná

The role of the Msx1 transcription factor in the intestinal epithelia and colorectal cancer

Ph.D. thesis

Supervisor: RNDr. Vladimír Kořínek, CSc.

Laboratory of Cell and Developmental Biology
Institute of Molecular Genetics of the ASCR, v.v.i

Prague, 2019

Announcement

I hereby declare that I wrote this doctoral thesis by myself and that I present my original research work. Contributions of others are clearly indicated and accompanied by reference to the literature and acknowledgements. The thesis has not been submitted, either in whole or in part, for the award of any other academic degree.

Prague, 1st March 2019

Mgr. Monika Šťastná

Acknowledgement

I would like to thank all people who made this thesis possible. My special thanks go to my supervisor Vladimír Kořínek for giving me the opportunity to work in his laboratory and for supervising my projects. I would also like to thank all of my colleagues for creating a wonderful work environment. In particular, I would like to thank Lucka Janečková who "took care of me" at the beginning of my Ph.D. studies and was always willing to help. Last but not least, my great thanks belong to my family and my husband for their generous support during my studies.

Contents

List of abbreviations

4-OHT	4-hydroxytamoxifen
ABHD12B	abhydrolase domain containing 12B
ACF	aberrant crypt foci
ACTNB	actin beta
ADAM10	a disintegrin and metalloproteinase domain-containing protein 10
aISCs	actively dividing intestinal stem cells
AKT	protein kinase B
AOM	azoxymethane
AP2	apetala 2
APC	adenomatous polyposis coli
Ascl2	achaete-scute complex homolog 2
Atoh1	atonal bHLH transcription factor 1
Axin	axis inhibition protein
bHLH	basic helix-loop-helix
BIO	(2'Z,3'E)-6-bromoindirubin-3′-oxime
Bmi1	B lymphoma Mo-MLV insertion region 1
BMP	bone morphogenetic protein
bp	base pair
BSA	bovine serum albumin
BTB	broad-complex, tramtrack, and bric-à-brac
βTrCP	beta-transducin repeat containing E3 ubiquitin protein ligase
BVES	blood vessel epicardial substance
CA-CRC	colitis-associated colorectal cancer
Cas9	CRISPR-associated protein 9
CBC	crypt base columnar
CD	Crohn's disease
cDNA	complementary DNA
CDK	cyclin dependent kinase
CDX2	caudal type homeobox 2
CHGA	chromogranin A
ChIP	chromatin immunoprecipitation
CK1α	casein kinase 1 alpha
cKO	conditional knock-out
CM	conditioned media
CRC	colorectal cancer
CRISPR	clustered regularly interspaced short palindromic repeats
CtBP	C-terminal-binding protein 1
CTNNB1	catenin (cadherin-associated protein), beta 1
CTSZ	cathepsin Z
DAB	diaminobenzidine
DAPI	4′,6-diamidine-2′-phenylindole dihydrochloride
DCS	deep crypt secretory

DEPDC7	DEP domain containing 7
Dkk	dickkopf
Dll1	delta-like 1
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulfoxide
DSS	dextran sodium sulfate
Dvl	dishevelled
EDTA	ethylene-diamine-tetra-acetic acid
EGF	epidermal growth factor
EGFP	enhanced green fluorescent protein
EMT	epithelial mesenchymal transition
ENTPD8	ectonucleoside triphosphate diphosphohydrolase 8
EphB	ephrin B
ERBB	erythroblastic oncogene B
Evi	evenness interrupted
FA	formaldehyde
FACS	fluorescence-activated cell sorting
FAP	familial adenomatous polyposis
FCCX	familial colorectal cancer type X
FISH	fluorescence *in situ* hybridization
FRT	flippase recognition target
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GFP	green fluorescent protein
Gli1	glioma-associated oncogene homolog 1
gRNA	guide RNA
GSK3β	glycogen synthase kinase 3 beta
Gy	gray
H3K4me3	trimethylation of lysine 4 on histone H3 protein subunit
H3K9me2	dimethylation of lysine 9 on histone H3 protein subunit
H3K27me3	trimethylation of lysine 27 on histone H3 protein subunit
H&E	hematoxilin and eosin
HDAC	histone deacetylase
HEK293	human embryonic kidney 293 cells
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HER2	human epidermal growth factor receptor 2
Hes1	hairy/enhancer of split 1
HGD	high-grade dysplasia
HIC1	hypermethylated in cancer 1
HNPCC	hereditary nonpolyposis colorectal cancer
Hopx	HOP homeobox
HP	hyperplastic polyps
HRP	horseradish peroxidase
HSPGs	heparan sulfate proteoglycans
HYP	hyperplasia

IBD	inflammatory bowel disease
IgG	immunoglobulin G
IMDM	Iscove's modified Dulbecco's medium
Int1	integration 1
IRES	internal ribosome entry site
ISCs	intestinal stem cells
kb	kilobase
kDa	kilodalton
KRAS	Kirsten rat sarcoma viral oncogene homolog
KRT	keratin
LEF	lymphoid enhancer-binding factor
LGD	low-grade dysplasia
Lgr5	leucine-rich repeat-containing G-protein coupled receptor 5
LRCs	label-retaining cells
Lrig1	leucine-rich repeats and immunoglobulin-like domains protein 1
LRP	low density lipoprotein receptor-related protein
Lyz1	lysozyme 1
MAPK	mitogen activated protein kinase
MCR	mutation cluster region
MDGA1	MAM domain containing glycosylphosphatidylinositol anchor 1
MDS	Miller-Dieker syndrome
Min	multiple intestinal neoplasia
Mmp7	matrix metalloproteinase 7
MMR	mismatch repair
MMTV	mouse mammary tumor virus
Msh	muscle segment homeobox
MSI	microsatellite instability
MSX1	msh homeobox 1
Mtus2	microtubule associated scaffold protein 2
Muc2	mucin 2
Mylk3	myosin light chain kinase 3
MyoD	myogenic differentiation antigen
mTert	mouse telomerase reverse transcriptase
NF-κB	nuclear factor kappa-light-chain-enhancer of activated B cells
Nkd1	naked cuticle 1
NSG	NOD/SCID/GAMMA
Olfm4	olfactomedin 4
PAS	periodic acid-Schiff
Pax	paired box
PBS	phosphate-buffered saline
PCNA	proliferating cell nuclear antigen
PCP	planar cell polarity
PFA	paraformaldehyde
PI	phosphatidylinositol

PI(3)P	phosphatidylinositol-3-phosphate
PI3K	phosphatidylinositol-3-kinase
PIK3CA	PI3K catalytic subunit α
PKCα	protein kinase C alpha
POZ	poxvirus and zinc finger
PRC2	polycomb repressive complex 2
PTEN	phosphatase and tensin homolog
qRT-PCR	quantitative real-time polymerase chain reaction
Rab	Ras-associated binding
RASL10B	RAS like family 10 member B
Reg4	regenerating gene 4
RFP	red fluorescent protein
rISCs	rarely dividing reserve ISCs
RNAi	RNA interference
Rspo	R-spondin
RT	room temperature
SFRP	secreted Frizzled-related protein
Shh	Sonic hedgehog
SI	sucrose isomaltase
Sirt1	sirtuin 1
Slc5a6	solute carrier family 5 member 6
SMAD	Smad family member
SNX	sorting nexin
SORBS2	sorbin and SH3 domain containing 2
Sox	sex-determining region Y (SRY)-related high-mobility-group box
SP5	SP5 transcription factor
Srt	sprinter
SSA	sessile serrated adenomas
STF	Super TOP-FLASH
STK32B	serine/threonine kinase 32B
Sulf	sulfatase
TA	transit amplifying
TALEN	transcription activator-like effector nuclease
TBP	TATA-box binding protein
TBS	Tris-buffered saline
TCF	T-cell factor
Tg	transgenic
TGFα	transforming growth factor α
TGN	*trans* Golgi network
TLR2	Toll-like receptor 2
TMEM47	transmembrane protein 47
Tnfrsf19	tumor necrosis factor receptor superfamily, member 19
TP53	tumor protein 53
Trpm	transient receptor potential cation channel, subfamily M

TSA	traditional serrated adenomas
Ttn	titin
UBB	ubiquitin B
UC	ulcerative colitis
VPS	vacuolar protein sorting
Wls	wntless
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Abstract

The Wnt signaling pathway represents the principal evolutionarily conserved signaling cascade found in all multicellular organisms. It plays a key role not only in many processes during embryogenesis, but also in maintaining tissue homeostasis and regeneration. By contrast, mutations in genes encoding components of the pathway often result in increased activation of Wnt signaling and underlie onset of many human diseases, particularly cancer.

The canonical Wnt signaling pathway is essential for proliferation and maintenance of the pluripotent state of intestinal stem cells and thus for homeostatic renewal of the intestinal epithelium. However, aberrant (hyper)activation of the Wnt signaling pathway is the initial step in development of intestinal neoplasia. Understanding the causes and identifying the consequences of the Wnt signaling hyperactivation is crucial for deciphering mechanisms leading to malignant transformation. Although the canonical Wnt signaling pathway has been the subject of scientific studies for several decades, all regulatory mechanisms and consequences of its hyperactivation have not been completely elucidated yet. During my PhD studies, I focused on understanding function(s) of some components and target genes of this signaling cascade. In this theses, results of my first author and one co-author publication are presented, which deal with two genes directly linked to Wnt signaling.

In my first-author publication, we studied the function of the msh homeobox 1 (MSX1) transcription factor in mouse and human intestines and tumors. We used a mouse model of the human disease Familial adenomatous polyposis (the $Apc^{+/min}$ mice) and mouse models harboring conditional knock-out alleles of the tumor suppressor gene adenomatous polyposis coli (*Apc*) and *Msx1*. These mice were intercrossed with mice expressing (regulated) Cre recombinase throughout the intestinal epithelium (*Villin-Cre* and *Villin-CreERT2*) or in the intestinal stem cells *(Lgr5-EGFP-IRES-CreERT2*), i.e. strains which enable spatiotemporal inactivation of the specific gene(s). We found that Msx1 is essential during formation of the so-called ectopic crypts, which are pouches of proliferating cells aberrantly occurring in the otherwise differentiated villous compartment after inactivation of the tumor suppressor gene adenomatous polyposis coli (*Apc*). Ectopic crypts have been described as a typical morphological feature of human serrated adenomas, which represent an aggressive type of intestinal polyps. Moreover, we suggest that Msx1 inactivation leads to a morphological conversion of intestinal tumors from tubular to villous adenomas, which

is in humans associated with more advanced stages along the path towards fully developed carcinoma and a worse survival prognosis. We also found out that Msx1 represents a robust marker of human colorectal carcinomas with the most elevated expression in the early stages of tumorigenesis. In the second publication, we described the role of a tumor suppressor hypermethylated in cancer 1 (*Hic1*) in mouse intestines. Using mice harboring conditional alleles of the *Hic1* gene and expressing Cre recombinase throughout the intestinal epithelium, we described that *Hic1* loss leads to increased numbers of differentiated intestinal epithelial cells and elevated levels of toll-like receptor 2 (*Tlr2*). Consequently, Tlr2 activates the NF-κB signaling pathway, which promotes intestinal tumorigenesis.

Key words: Wnt signaling, MSX1, HIC1, colorectal cancer, ectopic crypts

Abstrakt

Signální dráha Wnt představuje důležitou evolučně konzervovanou signální kaskádu, která se nachází u všech mnohobuněčných organismů. Signální dráha Wnt hraje klíčovou roli nejen v mnoha procesech během embryogeneze, ale také v udržování tkáňové homeostázy a při regeneraci. Oproti tomu mutace v genech kódujících komponenty signální dráhy Wnt často vedou k její zvýšené aktivaci a jsou příčinou vzniku mnoha lidských onemocnění, zejména rakoviny.

Kanonická signální dráha Wnt je nezbytná pro buněčnou proliferaci a udržování pluripotentního stavu střevních kmenových buněk, a tím pro homeostatickou obnovu střevního epitelu. Aberantní (hyper)aktivace signální dráhy Wnt představuje počáteční krok ve vývoji střevních neoplázií. Pochopení příčin a identifikace následků hyperaktivace signální dráhy Wnt je zásadní pro rozluštění mechanismů vedoucích k maligní transformaci. Ačkoli je kanonická signální dráha Wnt předmětem vědeckých studií již několik desetiletí, všechny regulační mechanismy a důsledky její hyperaktivace nebyly dosud zcela objasněny. Během mého doktorského studia jsem se zaměřila na pochopení funkcí některých komponent a cílových genů této signalizační kaskády. V této práci jsou prezentovány výsledky mé prvoautorské a jedné spoluautorské publikace, které se zabývají dvěma geny přímo spojenými se signální dráhou Wnt.

V rámci mé prvoautorské publikace jsme studovali funkci transkripčního faktoru msh homeobox 1 (MSX1) v myší a lidské střevní tkáni a nádorech. Použili jsme myší model lidského onemocnění Familiální adenomatózní polypóza, tedy myši se zkrácenou alelou tumor supresorového genu adenomatous polyposis coli (*Apc*), které v dospělosti vyvinou četné střevní polypy (myši kmene $Apc^{+/min}$). Dále jsme použili myši, které obsahovaly podmíněné alely genů *Apc* a *Msx1*. Tyto myši byly kříženy s myšími kmeny, které exprimovaly (regulovanou) Cre rekombinázu v celém střevním epitelu (*Villin-Cre* a *Villin-CreERT2*) nebo ve střevních kmenových buňkách (*Lgr5-EGFP-IRES-CreERT2*) - tedy s kmeny umožňujícími časově i prostorově regulovanou inaktivaci specifických genů. Zjistili jsme, že protein Msxl je nezbytný při tvorbě takzvaných ektopických krypt, což jsou kapsovité útvary tvořené proliferujícími buňkami, jež se po inaktivaci genu *Apc* aberantně vyskytují v jinak diferencovaném vilovém kompartmentu. Ektopické krypty byly popsány jako typický morfologický rys lidských adenomů se „zoubkovaným“ uspořádáním epitelu krypt („serrated“ adenoma), které představují agresivní typ střevních polypů. Kromě toho se domníváme, že inaktivace genu *Msx1* vedla k morfologické konverzi střevních nádorů z

tubulárních na vilózní adenomy, což je u lidí spojeno s pokročilejšími stádii kolorektálních karcinomů a horší prognózou přežití. Zjistili jsme také, že *MSX1* představuje robustní marker lidských kolorektálních karcinomů s nejvyšším výskytem v počátečních stádiích tumorigeneze. Ve druhé publikaci jsme popsali úlohu nádorového supresoru hypermethylated in cancer 1 (*Hic1*) v myší střevní tkáni. S využitím myší, které obsahují podmíněné alely genu *Hicl* a exprimují Cre rekombinázu v celém střevním epitelu, jsme popsali, že ztráta Hicl vede ke zvýšení počtu diferencovaných střevních epiteliálních buněk a zvýšené hladině mRNA toll like receptor 2 (*Tlr2*). Protein Tlr2 následně aktivuje signální dráhu NF-kB, která podporuje střevní tumorigenezi.

klíčová slova: signální dráha Wnt, MSX1, HIC1, rakovina tlustého střeva, ektopické krypty

1 Overview of the literature

1.1 The Wnt signaling pathway

The Wnt signaling pathway has been the subject of many scientific studies since 1982, when Roel Nusse and Harold Varmus discovered proto-oncogene integration 1 (*Int1*), which was activated in mouse breast tumors upon infection with mouse mammary tumor virus (MMTV). *Int1* was the first described member of the Wnt gene family. The family named was derived as an acronym of two gene names, *wingless*, the *D. melanogaster Int1* paralog, and *Int1;* the letter gene was renamed as *Wnt1*. Over more than three decades, many important discoveries have been accomplished in the Wnt signaling field across various areas of biology ranging from embryonic development to homeostasis in adult tissues and numerous human diseases, including cancer. Although a great number of Wnt signaling pathway components and mechanisms of its regulation have been already described, there is still much to be discovered in this area of cellular signaling research.

Wnt proteins are extracellular signaling molecules that act as morphogens in regulation of various processes during embryonic development, such as cell fate determination, cell polarity, proliferation, and migration (reviewed in[35, 407]). In adults, Wnt signaling is essential for tissue homeostasis and regenerative processes following injury (reviewed in[195, 403]). Disruption or misregulation of Wnt signaling underlies development of human malignancies and is involved in various degenerative diseases (reviewed in[35, 109]).

Multiple *Wnt* genes have been identified in most metazoans (reviewed in[113]). In mammals, 19 genes encoding different Wnt proteins have been discovered so far (reviewed in[217]). Various Wnt ligands in combination with their receptors from the Frizzled family and co-receptors from the low density lipoprotein receptor-related protein (LRP) family activate several different branches of the Wnt signaling pathway: the β-catenin-dependent so-called canonical Wnt signaling pathway, planar cell polarity (PCP) pathway, and Wnt/calcium pathway (reviewed in[387, 407]). As there is a distinctive cross-talk between these pathways, the more accurate idea of Wnt signaling is more like a network than individual independent pathways (reviewed in[375]).

The research topic I dealt with during my PhD studies concerns the canonical Wnt/β-catenin pathway, therefore other parts of the Wnt signaling network are not discussed. Information on the non-canonical, i.e. β-catenin-independent, Wnt signaling pathways may be found in numerous (recent) reviews[50, 119, 353].

1.1.1 Situation in the Wnt-producing cell

Wnt proteins synthesized in Wnt-producing cells undergo several posttranslational modifications essential for their delivery to the extracellular space and proper binding to their receptors on target cells. All Wnt proteins consist of approximately 350 to 400 amino acids including 23 to 24 conserved cystein residues, which play an important role when folding their globular secondary structure (reviewed in[217, 242]). Mammalian Wnt proteins contain a signal peptide that enables their targeting to the endoplasmic reticulum where multiple oligosaccharides are appended to the asparagine residues. Interestingly, in polarized epithelial cells such as of the small intestine, glycosylation processing determines the secretory route of Wnt proteins to the apical or basolateral region of cell membrane[413]. The attachment of oligosaccharides presumably facilitates subsequent acylation[171] and binding to extracellular heparan sulfate proteoglycans (HSPGs)[8]. Wnts are further modified by palmitoleoylation at a conserved serine residue (Ser209 in murine Wnt-3a) by membrane-bound O-acyltransferase porcupine[155] and palmitoylation at a conserved cystein residue (Cys77 in murine Wnt-3a)[355]. These lipid modifications are indispensable for proper extracellular transport of Wnt proteins and their signaling activities[171, 405]. Once modified, Wnt proteins are sorted by p24 receptor[32, 270] to coat protein II vesicles and carried to the Golgi apparatus, where they are bound to wntless (Wls)/evenness interrupted (Evi)/sprinter (Srt) transmembrane chaperon[11, 92], and subsequently, transported within Rab8-positive vesicles to the plasma membrane[47]. Upon release of its cargo, Wls is internalized by clathrine-mediated endocytosis from the plasma membrane to Rab5-positive early endosomes and sorted back to the trans-Golgi network for another rounds of Wnt secretion[84] (Figure 1). This retrograde transport is mediated by retromer[20, 78, 251, 269, 417], a protein complex evolutionarily conserved from yeast to mammals[99]. The core assembly of retromer is a trimer of vacuolar protein sorting (VPS) proteins (Vps26/Vps29/Vps35)[314] that recognizes cargo proteins. Associated sorting nexin (SNX) heterodimer[315] (Vps5/Vps15 in yeast) binds to phosphatidylinositol-3-phosphate [PI(3)P] molecules[46], bent the plasma membrane[38] and consequently enable recruitment of the Vps trimer. Retromer binding to plasma membrane and thus recyclation of Wls is regulated by PI kinases[33] and PI(3)P phosphatases from myotubularin family[332] that modulate the (phospho)lipid composition of the plasma membrane and therefore influence its affinity to retromer. When the retrograde path is impaired, Wls molecules are targeted from early

endosomes to Rab7/Rab9-positive late endosomes/lysosomes where they are degraded. Consequently, Wnt proteins are retained within the cell and the signaling is silenced.

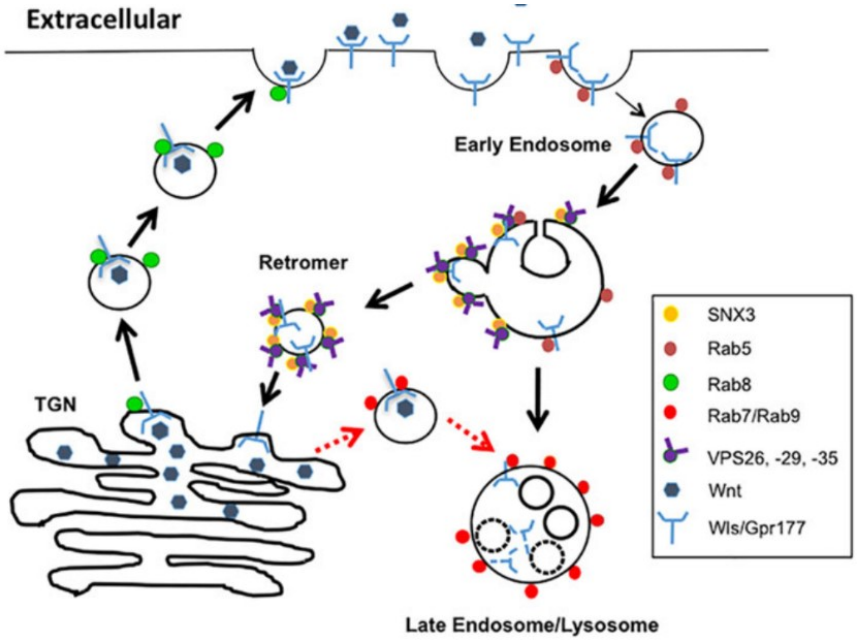


Figure 1 | Model of Wnt secretion and Wls recyclation. Wnt proteins in the *trans* Golgi network (TGN) associate with their receptor Wntless (Wls) and are delivered to the plasma membrane in Rab8-positive vesicles. After Wnt release, Wls is endocytosed to Rab-5 positive early endosomes and either recycled to TGN via retromer-mediated transport route or transported to Rab7/Rab9-positive late endosomes for degradation (adopted from[47]).

1.1.2 Wnt ligands transport in the extracellular space

Wnt proteins are known to form anteroposterior gradients during embryonic development[160, 238, 241]; however, they can not diffuse freely in the extracellular space due to their highly hydrophobic nature and therefore remain closely attached to cell surface or extracellular matrix molecules. Various theories postulate possible mechanisms of how Wnt proteins reach their final destinations. In short-range signaling, Wnt ligands are transferred directly from the Wnt-producing to the Wnt-receiving cell that are in close vicinity, such as Paneth and stem cells within the small intestinal epithelium[66]. Another way how Wnt proteins can move on short distances is via association with HSPGs of the glypican family[8], cell surface molecules that are components of the extracellular matrix (reviewed in[191, 307]). The long-range signaling may be realized in several mechanisms that enable movement of Wnt molecules and thus spreading of the signal through extracellular environment. Numerous studies described that Wnt proteins are released from cells inside various vesicles such as argosomes[93], exosomes[96], exosome-like vesicles[18, 174], or

lipoprotein particles[237, 252]. Wnts can also be transported via signaling filopodia (reviewed in[346]), dynamic cytoplasmic extensions that connect distant cells. Another mechanism of the Wnt transport to distant places is on migrating cells. For example, in the chick embryo, Wnt proteins produced by neurepithelial cells of the dorsal neural tube are loaded to migrating neural crest cells that physically transport the cargo to target cells of the dorsomedial lip, a medial compartment of dermomyotome[318]. Intriguingly, a recent study on *C. elegans* shows directly for the first time that Wnt ligands are capable of free dispersal in the extracellular space to establish long-range Wnt gradient[253].

Upon release from Wnt-producing cell, Wnt ligands may be recognized and bound by proteins that modulate the signaling amplitude. A group of natural Wnt antagonists inhibit the signaling by direct binding to Wnt ligands and preventing their engagement with the Frizzled and LRP receptor complexes, e.g. secreted Frizzled-related proteins (SFRP)[187, 190, 397], Wnt inhibitory factor-1[117], and Cerberus[267]. Alternatively, Wnt antagonists bind directly to the Frizzled/LRP receptors, thus blocking their binding sites for Wnt ligands. To this group of inhibitors belong proteins from Dickkopf (Dkk) family Dkk1, Dkk2, and Dkk4 that bind the LRP5/6 co-receptors and together with the Kremen receptors trigger their internalization from the plasma membrane, thus preventing formation of the Frizzled/LRP complexes[67, 179, 205-207, 317]. LRP5/6 co-receptors may also be bound by other secreted Wnt inhibitors, such as Wnt modulator in surface ectoderm[142] and Sclerostin[188]. Wnt ligands may as well get engaged with positive regulators, e.g. an extracellular matrix protein Periostin that recruits Wnts to mouse metastatic breast cancer stem cells, therefore increasing the level of Wnt signaling leading to metastatic colonization of the lungs[204]. Wnt signaling may be also promoted by endosulfatases Sulf1 and Sulf2 which desulfates heparin-sulfate chains on glypicans and thus loosen HSPGs linkage to Wnts, releasing the ligands for binding to their receptors[58]. Interestingly, Sulf1 expression is activated by Wnt signaling itself and therefore contributes as a feedback loop to establishment of the Wnt gradient[169].

1.1.3 Situation in the Wnt-receiving cell

The so-called canonical Wnt signaling pathway is the best studied part of the Wnt signaling network. It differs from other branches of the Wnt signaling network by the fact that its activity is regulated by degradation or stabilization of β-catenin protein, therefore it is also often referred to as the Wnt/β-catenin pathway (Figure 2). The β-catenin protein is

divided within the cell into two separate pools with distinct functions. Firstly, the β-catenin pool exerting structural functions is engaged in cell-cell junctions where it associates with E-cadherins[262]. Secondly, free cytosolic β-catenin is under certain conditions translocated to the cell nucleus where it interacts with transcription factors from T-cell factor/lymphoid enhancer-binding factor (TCF/LEF) family and accomplishes its signaling functions by modulating expression of Wnt signaling target genes[19, 224].

In a situation when the Wnt ligand is not present, β-catenin is marked for degradation by a protein complex called the β-catenin destruction complex. The complex contains casein kinase 1 alpha (CK1α)[192] and glycogen synthase kinase 3 beta (GSK3β)[297] which are responsible for β-catenin phosphorylation. Interaction between the kinases and β-catenin is enabled by scaffold proteins axis inhibition protein[135] (Axin) 1/2 and adenomatous polyposis coli (APC)[104], which are both tumor suppressors. CKIα and GSK3β mediated phosphorylations promote β-catenin ubiquitination by beta-transducin repeat containing E3 ubiquitin protein ligase (βTrCP) and its subsequent proteasomal degradation[1, 168]. This mechanism allows cells to maintain low levels of β-catenin in the cytoplasm and therefore low expression of Wnt target genes, as the effector transcription factors from the TCF/LEF family are bound by a transcription corepressor Groucho[42, 293] (Figure 2, left).

When Wnt signaling is active, Wnt ligands bind to cell surface receptor complex composed of the Frizzled receptor and coreceptor LRP5/6. This interaction induces a cascade of intracellular events that lead to phosphorylation of the adaptor protein Dishevelled (Dvl) and its recruitment to the Frizzled/LRP5/6 complex. Subsequently, Axin protein is recruited by Dvl to the cell membrane which results in breakup of the β-catenin destruction complex. Beta-catenin then translocates to the cell nucleus where it associates with transcription factors from the TCF/LEF family and multiple co-activators and initiates transcription of the Wnt target genes (Figure 2, middle; reviewed in[200]), e.g. achaete-scute complex homolog 2 (*Ascl2*)[153], *Axin2*[152, 198, 414], *cyclin D1*[325, 366], *cMyc*[107], naked cuticle 1 (*Nkd1*)[425], olfactomedin 4 (*Olfm4*)[194], and tumor necrosis factor receptor superfamily, member 19 (*Tnfrsf19*, alternative name *Troy*)[64, 281].

In the absence of the Wnt stimulus, the pathway may be aberrantly activated as a result of mutations in genes encoding key components of the pathway. The mutations mostly include inactivating mutations of the tumor suppressor *APC* (Figure 2, right).

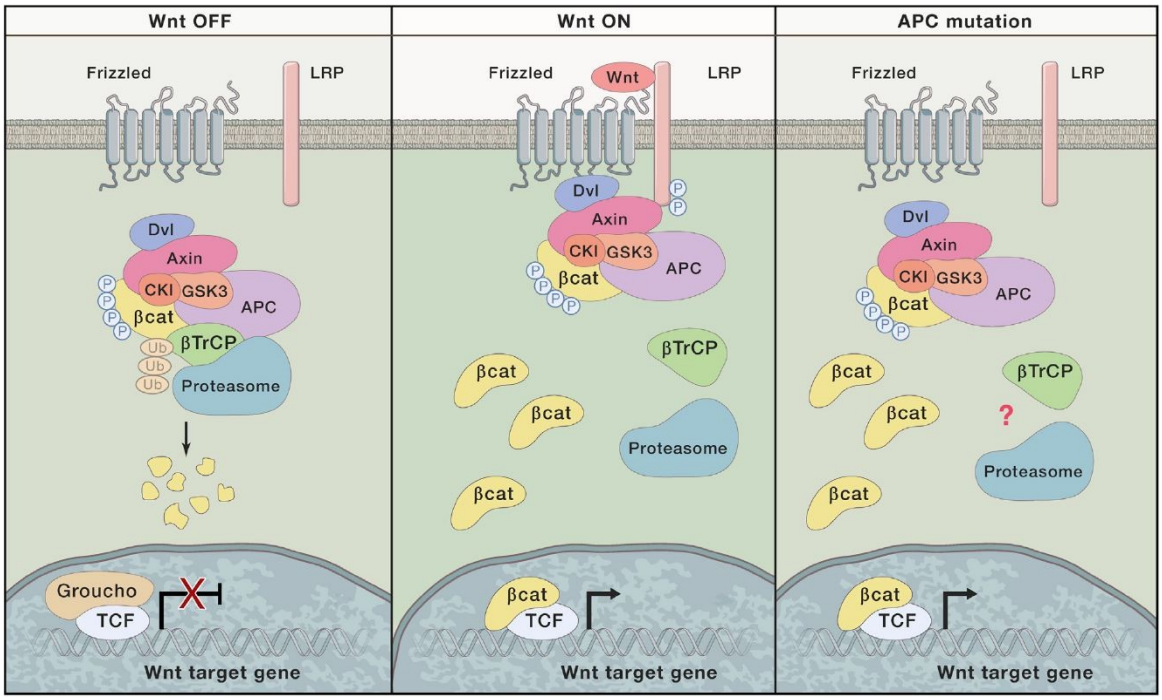


Figure 2 | The canonical Wnt signaling pathway.
Left, in the absence of Wnt stimulus, the so-called β-catenin destruction complex composed of adenomatous polyposis coli (APC), axis inhibition (Axin), casein kinase 1 alpha (CK1α), glycogen synthase kinase 3 beta (GSK3β), and beta-transducin repeat containing E3 ubiquitin protein ligase (βTrCP) enables phosphorylation and subsequent ubiquitination of β-catenin (βcat) leading to its degradation in proteasome. T-cell factor/lymphoid enhancer-binding factor (TCF/LEF) transcription co-factors are bound by the transcription repressor Groucho in an inactive state and therefore the Wnt target genes transcription is turned off. Middle, association of Wnt ligand with its receptor Frizzled and co-receptor low density lipoprotein receptor-related protein (LRP) induces LRP phosphorylation that leads to Axin recruitment to the cell membrane followed by disassembly of the destruction complex. Stabilized β-catenin then translocates to the cell nucleus and together with its transcription co-factors TCF/LEF activates transcription of Wnt target genes. Right, inactivating mutations in the *APC* tumor suppressor gene result in production of the truncated protein that in most cases has no longer the ability to scaffold the destruction complex, therefore β-catenin is not degraded and may enter the nucleus. In such cases, the Wnt pathway is activated even without presence of the Wnt ligand (adopted from[243]).

1.2 The gut

Mammalian bowel is a long tube connecting the stomach and the rectum. The organ consists of the small and large intestine. The main function of the small intestine is to absorb water and nutrients from food, but it also acts as an important constituent of the immune system. Duodenum is the initial segment of the small intestine where the food from the stomach comes and into which bile and pancreatic ducts deliver essential enzymes for food digestion. The following and the longest section of the small intestine is jejunum, which is terminated by ileum that enters caecum. From the caecum comes out the colon which is responsible for final absorption of nutrients and water, synthesis of vitamins and elimination of feces through rectum out of the body.

The intestinal wall consists of four layers of specialized tissue. Serosa is the outermost layer of connective tissue interwoven with blood and lymphatic vessels and nerves. The underlying muscularis propria consisting of two smooth muscle layers (outer longitudinal and inner circular layer) induces rhythmic contractions that generate the peristaltic gut movement essential for proper locomotion of the chymus. The submucosa is a dense irregular connective tissue layer undearneath; in the duodenum it contains Brunner's glands that produce alkaline mucus necessary for neutralization of the acidic chymus coming from the stomach. The innermost layer, the mucosa, is further divided into three layers: a thin smooth muscle layer (muscularis mucosae), connective tissue layer (lamina propria), and finally the single cell layer of epithelial cells lining the gut lumen. Being completely renewed every 3-5 days, the small intestinal epithelium exhibits one of the highest turnovers from all human tissues and therefore represents a suitable object for studying mechanisms of tissue regeneration and adult stem cells, as various phenotypic changes in the tissue occur already within couple of days after the damage or genetic manipulation (reviewed in[208]).

1.2.1 Architecture of the intestinal epithelium

The single layer epithelium that lines the lumen of both the small intestine and the colon permeates into the underlying mesenchyme to form gland-like pits called crypts which serve as niche for intestinal stem cells (ISCs). The inner surface of the small intestinal mucosa is increased by circular folds (*plica circulare*) covered with finger-like protrusions called villi which are composed of fully matured cells responsible for digestion- and resorption-associated functions of the tissue. Contrary, the inner surface of the colon is

flat and matured cell types occupy upper parts of the crypts. Every 24 hours, the ISCs undergo a symmetrical division and either retain their stemness or exit the crypt base and become transit amplifying (TA) cells[339]. TA cells are committed progenitors that migrate through the upper region of the crypt where they undergo up to six rounds of cell divisions, commencing differentiation at around the third generation (reviewed in[208]). Once leaving the crypt-villus border, TA cells terminally differentiate into one of the absorptive or secretory cell lineages and migrate towards the top of the villus where they are "squeezed out" into the gut lumen. The process is closely interconnected with programmed cell death, although it is still an open question whether apoptosis is the cause or effect of epithelial cells shedding (reviewed in[260]).

Absorptive enterocytes with microvilli protruding from their apical membrane that further increase the epithelial surface for nutrient and water uptake represent the most numerous cells of the small intestinal epithelium. The most abundant secretory cells are goblet cells, which represent approximately 4 % of epithelial cells in the duodenum and exhibit increasing abundance along the anterior-posterior axis into approximately 16 % in the distal colon (reviewed in[157]). Goblet cells secret mucin, the major component of mucus layer which is an important protective barrier between the epithelial cells and intestinal contents. Enteroendocrine cells producing peptide hormones represent a minor population (<1%)[347] of secretory cells. Paneth cells are the only known population of differentiated cells that reside at the crypt base. They persist in the crypt for 6-8 weeks[139], contribute to stem cell niche maintenance[310] and produce antimicrobial peptides, such as lysozyme 1 (Lyz1)[261] and α-defensins (alternative name cryptdins)[7, 401]. Two minor cell populations of recently discovered cells seem to be predominantly active within the immune response to pathogens, these are tuft cells producing opioids and cyclooxygenase enzymes[89] and M-cells that uptake intraluminal antigens and transport them to lymphocytes in the Peyer's patches lying beneath the epithelial layer[53]. Finally, enterocytes-resembling brush cells are abundant in the epithelium adjacent to the Peyer's patches; their contacts with axons suggest chemosensory functions[87, 229]. Although stem cells were discovered also in colon crypts[12], Paneth cells were not. Nevertheless, colon crypts also contain supporting cells. Recent studies identified a group of regenerating gene 4 (Reg4)-positive deep crypt secretory (DCS) cells[308] as a subpopulation of cKit/CD117-positive goblet cells[296]. These cells seem to fulfill the niche-maintainig functions within the colonic crypt base.

1.2.2 Pathways regulating intestinal homeostasis

The equilibrium between cell proliferation and differentiation is governed by a tightly regulated network of various signaling cascades that establish gradients of signaling molecules along the crypt-villus axis (Figure 3). The major driving force behind intestinal crypt maintenance is the Wnt/β-catenin signaling that synergistically with the Notch signaling pathway sustains the proliferative and undifferentiated status of stem and progenitor cells. Additionally, proliferation of stem and progenitor cells is supported by epidermal growth factor (EGF) signaling. Conversely, differentiation is inhibited within the crypt base by production of bone morphogenetic protein (BMP) antagonists. On the other hand, BMP signaling together with Hedgehog pathway promote cell differentiation in the villi. The signaling system of Ephrin B (EphB) receptors and ephrin-B ligands, expressed in opposing gradients, contributes to segregation of differentiated cells. Deregulation of these signaling pathways often disturbs cellular composition of the epithelium and eventually may result in neoplasia formation (reviewed in[176, 177, 226]).

The increasing gradient of Wnt ligands towards the crypt base ensures proliferative and undifferentiated status of stem and progenitor cells, and, moreover, contributes to proper differentiation of Paneth cells[4, 233]. Disruption of Wnt signaling by expression of Wnt inhibitors such as Dkk1 or by removal of Tcf4 or β-catenin leads to loss of stem cells which is accompanied by degeneration of the epithelial architecture[138, 173, 180]. In contrast, aberrant activation of Wnt signaling, mostly triggered by inactivating mutations in the *Apc* gene, causes hyperproliferation of intestinal stem cells followed by enlargement of the crypt compartment[305] leading to neoplastic formation and tumor development in the small intestine and colon.

The Wnt signaling pathway is sophistically modulated in order to control stem cells renewal. The cooperation of the Hedgehog and BMP signaling pathways counterbalance Wnt activity to prevent stem cells hyperproliferation and facilitate differentiation of their descendants[108, 379]. In brief, Sonic hedgehog (Shh) and Indian hedgehog ligands produced by TA cells bind to their receptor Patched on the surface of mesenchymal cells, which induces production of BMPs. Soluble BMP molecules bind their serine/threonine kinase receptors on mesenchymal and epithelial cells that transduce the signal to the nucleus through the Smad family member (SMAD)1/5/8 and SMAD4 transcription (co)factors or via other signaling cascades, e.g. phosphatidylinositol-3-kinase (PI3K)/AKT, p38 or c-Jun N-terminal kinase (reviewed in[398]). SMAD-mediated transcription repression affects

expression of stem cell signature genes and thus restricts stem cells expansion[280]. While the progenitor cells are exposed to increasing concentration of pro-differentiation BMP signals during their migration towards the crypt-villus border, stem cells residing at the bottom of the crypt are exposed to high concentrations of BMP inhibitors. In the small intestine, mesenchymal cells produce a BMP antagonist Noggin that sequesters BMP ligands and thus prevents their binding to BMP receptors (reviewed in[103]). Colonic stem cells are protected by BMP antagonists gremlin 1/2 and chordin-like 1 which are produced by myofibroblasts and smooth muscle cells surrounding the colon crypts[175].

The decreasing gradient of Wnt ligands towards the top of the villus induces a reverse gradient of EphB receptors and ephrin-B transmembrane ligands expression. This signaling system governs compartmentalization along the crypt-villus axis and proper localization of epithelial cells by modulating their adhesion through E-cadherin containing cell junctions[16] (reviewed in[44]). High activity of Wnt signaling in cells residing at the crypt base (stem and Paneth cells) induces expression of EphB3 receptor and simultaneously inhibits expression of ephrin-B1 ligand; EphB2 receptor is expressed in all proliferating crypt cells with a decreasing intensity from the crypt bottom to the crypt-villus border[16]. When the progenitor cells migrating towards the crypt orifice escape the zone of active Wnt signaling, repression of ephrin-B1 ligand weakens. At the interface between populations expressing EphB receptor and ephrin-B1 ligand, a locally activated metalloproteinase ADAM10 cleaves E-cadherin, which leads to asymmetric distribution of EphB receptor and, consequently, altered affinity between the two cell populations. Therefore the lack of EphB2 expression in Paneth cells prevents them from escaping the crypt compartment and high EphB3 expression level facilitates their adhesion to cells in the crypt bottom[342].

The Wnt and Notch signaling pathways work synergistically to preserve the undifferentiated and proliferating stem and progenitor cells in the crypt compartment[380]. In stem cells, the Notch signaling is activated via interaction of redundant Notch1 and Notch2 receptors with delta-like 1 (Dll1) and Dll4 transmembrane ligands present on the surface of Paneth cells[79, 264]. The signal is then transmitted by activation of hairy/enhancer of split 1 (Hes1) transcription factor expression, which inhibits transcription of cyclin dependent kinase (CDK) inhibitors $p27^{Kip1}$ and $p57^{Kip2}$ [286]. Notch signaling is also important during decision making between adoption of absorptive or secretory cell fate. Hes1 inhibits atonal homologue 1 (Atoh1), a transcription factor that promotes adoption of secretory cell fate, and thus induces progenitor cells differentiation into absorptive enterocytes (reviewed in [330]).

Signaling initiated by ligands that belong to the EGF family or a closely related ErbB/HER/Neu family is transmitted through their tyrosine kinase receptors that activate several signaling pathways responsible for cell proliferation and survival, such as mitogen activated protein kinase (MAPK) cascade. At the same time, intestinal stem cells regulate the extent of niche expansion by elevated expression of transmembrane leucine-rich repeats and immunoglobulin-like domains protein 1 (Lrig1)[247, 408] which attenuates the EGFR/ErbB signaling in a negative feedback manner[98] and, simultaneously, promotes the anti-proliferative BMP signaling[247]. In addition to proliferation-promoting activity, the EGFR signaling is also important in inhibiting apoptosis of stem and progenitor cells[351].

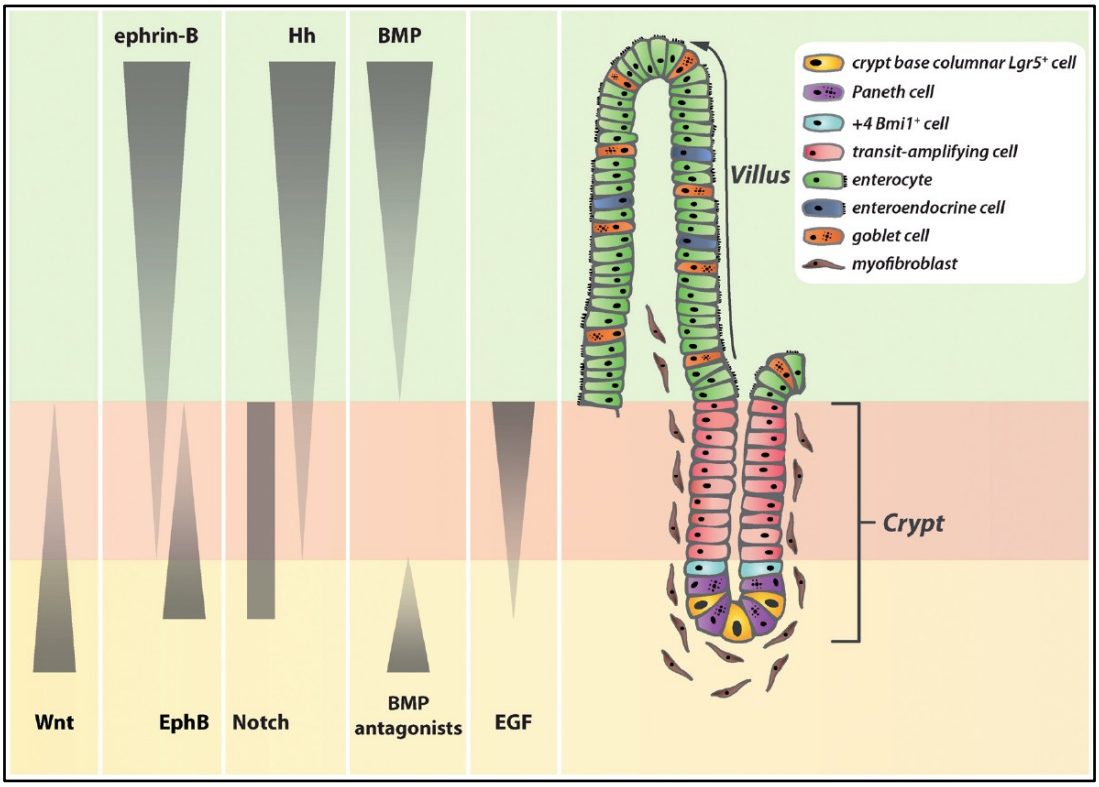


Figure 3 | Signaling cascades regulating homeostasis of the intestinal epithelium.
At the crypt bottom reside fast cycling Leucine-rich repeat-containing G-protein coupled receptor 5 (Lgr5)-positive stem cells among post-mitotic Paneth cells. Four cell diameters from the crypt bottom is a quiescent population of reserve, $Bmi1^+$ inestinal stem cells. Stem cells generate transit amplifying (TA) cells, which are progenitors of differentiated lineages. The balance between proliferation and differentiation is regulated by a complex network of signaling cascades. The Wnt and Noth pathways synergistically maintain proliferating state of stem and TA cells and have important functions in lineage specification. An oposing gradient of EphB receptors and ephrin-B ligands enables spatial segregation of crypt compartments. BMP and Hedgehog pathways activity on villi restrain proliferation and promote differentiation. BMP signals at the crypt bottom are revoked by BMP antagonists of mesenchymal origin. EGF-dependent signaling induces proliferation and inhibits apoptosis in TA cells (adopted from[176]).

1.2.3 Intestinal stem cells

As already indicated, the intestinal epithelium is very dynamic and the turnover of differentiated cells is approximately 3 to 5 days. To ensure such fast renewal, very rapidly dividing stem cells in the crypt base continuously refill the compartment of proliferating TA (progenitor) cells which then give rise to differentiated cell lineages. Stem cells of the small intestine were first recognized by physician Joseph Paneth, who already in 1887 depicted in his hand drawing large granular cells at the bottom of the small intestinal crypt (later named Paneth cells) and morphologically very different cells incised among them (Figure 4A). Several decades later, these cells were described as "genuine" ISCs (reviewed in[45]).

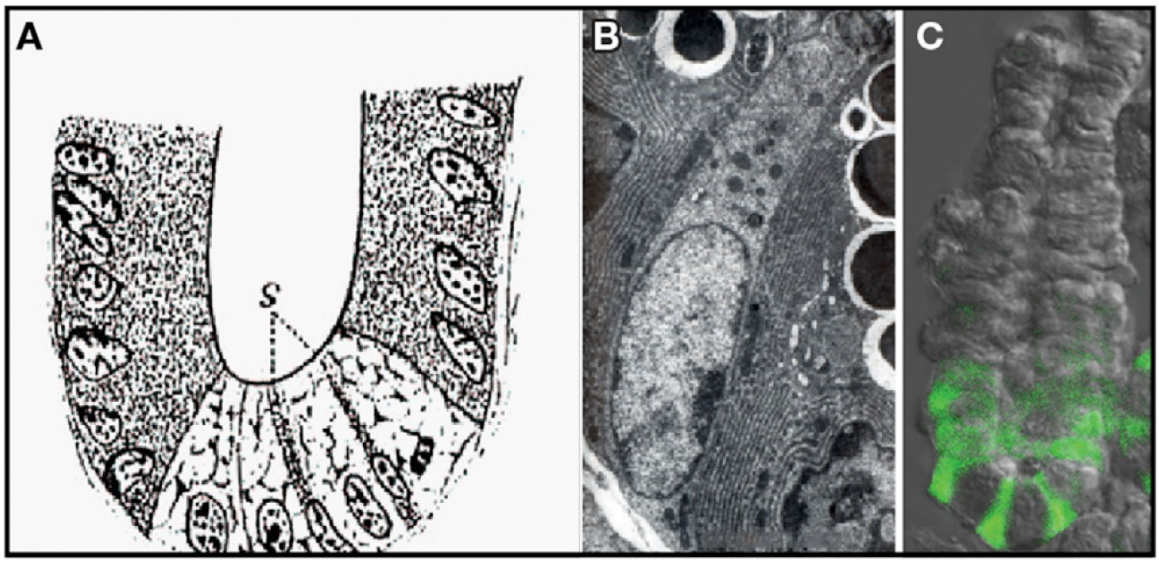


Figure 4 | Visualization of the Paneth cells and intestinal stem cells.
(A) Joseph Paneth's hand drawing of the small intestinal crypt showing the Paneth cells (large white cells) and prospective stem cells (thin dark cells; indicated by dotted line and letter "s"); (B) First image of the intestinal stem cell between two Paneth cells (with distinct lysosyme-containing granules) obtained by electron microscope; (C) Confocal image of the crypt isolated from *Lgr5-EGFP-IRES-CreERT2* mouse, stem cells endogenously express green fluorescent protein (GFP; adopted from[45]).

1.2.3.1 Intestinal stem cell populations

Fast-cycling stem cells / Crypt base columnar (CBC) cells

The first mention of rapidly dividing cells located at the bottom of small intestinal crypts dates back to 1974 when Cheng and Leblond showed that all cell lineages of the intestinal epithelium originate from a single source[130]; moreover, they displayed for the first time intestinal stem cells [termed crypt base columnar (CBC) cells] by electron

microscopy (Figure 4B)[129]. In the following years, several experiments using chemical mutagenesis confirmed that these cells are indeed stem cells and, moreover, the experiments revealed the clonal character of the crypts[25]. With more advanced methods of genetic manipulations, more elegant ways of lineage tracing were implemented. In 2007, Barker and colleagues identified the Wnt target gene leucine-rich repeat-containing G-protein coupled receptor 5 (Lgr5)[376], a transmembrane receptor of small molecules R-spondins (Rspo) that enhance Wnt signaling[52], as a marker of rapidly cycling stem cells residing at the crypt base in both small intestine and colon[12]. Introduction of a knock-in allele *Lgr5-EGFP-IRES-CreERT2* into the mouse genome enabled direct visualization of CBC cells (Figure 4C), as the green fluorescent protein (GFP) is expressed directly from the *Lgr5* locus. Since *Lgr5-EGFP-IRES-CreERT2* mice produce inducible CreERT2 enzyme after crossing to *ROSA26-LacZ* reporter animals, lineage tracing experiments of cells originating from $Lgr5^+$ stem cells were performed. Within five days after CreERT2 recombinase activation, $Lgr5^+$ cells generated all differentiated cell lineages. The LacZ signal was observed as ribbons of blue cells running from the crypt bases to top of the villi; the signal persisted in the tissue for more than 60 days (many blue ribbons remained for life-long), confirming that the $Lgr5^+$ cells represent small intestinal and colonic stem cells[12]. Later on, Sato and co-workers performed fluorescence-activated cell sorting (FACS) of epithelial cells obtained from the *Lgr5-EGFP-IRES-CreERT2* mouse intestines and succeeded to establish *in vitro* culture of "organoids", i.e. self-organized 3D structures of epithelial cells resembling normal intestine, that originated from individual $Lgr5^+$ (GFP^{high}) cells[309]. Subsequent gene expression profiling of $Lgr5^+$ cells enabled identification of ISCs gene signature and led to the discovery of many other ISC markers[232, 377]. One of them was the Wnt target gene *Ascl2* which is essential for ISCs fate, as its depletion from the tissue results in loss of $Lgr5^+$ stem cells within a few days[377]. Ascl2 is a transcription factor that is regulated by Wnt/Rspo signaling and synergistically with β-catenin/Tcf4 complexes activates expression of stem cell-specific genes[327]. Levels of *Olfm4* and *Troy* were also elevated in mouse small intestinal stem cells; however, these two genes are not expressed in mouse colon[377]. Following lineage tracing experiments verified Troy expression in fast cycling ISCs and revealed its association with Lgr5. Interestingly, depletion of Troy from small intestinal organoids resulted in a reduced need for the Wnt agonist Rspo, indicating that Troy acts as a negative regulator of Wnt/Rspo signaling[64]. Last but not least, EGF/ErbB inhibitor Lrig1 was abundant in ISCs and LacZ-mediated labeling of Lrig1-expressing progeny showed ribbons of cells running along the crypt-

villus axis[275, 408]. However, expression profiling of colonic stem cells revealed remarkable differences between $Lgr5^+$ and $Lrig1^+$ colonic stem cells, suggesting existence of two stem cell populations, a proliferating and a quiescent one[275], which will be described in more detail further.

Slow-cycling stem cells / +4 quiescent stem cells

Already in 1978, Potten and colleagues discovered during their studies of DNA segregation in epithelial cells a small population of DNA label-retaining cells (LRCs) that reside in the so-called +4 position (4 cells above the crypt base) and suggested that these cells are mitotically quiescent[273]. Later on, several research groups independently identified a population of cells inhabiting the +4 position, which are able to regenerate the crypt compartment upon depletion of $Lgr5^+$ stem cells by irradiation[282] or through transgenic expression of diphteria toxin receptor[369]. These slow-cycling cells are insensitive to perturbations of Wnt signaling, resistant to irradiation, and only weakly participate in homeostatic epithelial regeneration. However, in case of irradiation-induced injury, these cells start to proliferate and replenish the irradiation-sensitive $Lgr5^+$ population[415]. Many research groups sought to identify specific markers of the +4 cells. The first discovered marker was a polycomb protein B lymphoma Mo-MLV insertion region 1 (Bmi1)[302]. FACS-sorted $Bmi1^+$ cells were able to establish self-renewing organoid culture with all differentiated cell lineages, confirming their function as ISCs[415]. However, lineage tracing experiments resembled results obtained from the Lgr5 mouse strain, suggesting Bmi1 expression also in CBC cells[302]. Mouse telomerase reverse transcriptase (mTert) was proposed to be another +4 cells marker, as it was specifically expressed in injury-resistant cell population with the ability to reconstitute all differentiated intestinal cell lineages upon injury[225]. Protein HOP homeobox (Hopx) is also abundant in +4 cells; *Hopx-LacZ* reporter mouse showed long-persisting labelling and $Hopx^+$ cells gave rise to full-fledged organoids. Interestingly, $Hopx^+$ and $Lgr5^+$ cells were able to interconvert, supporting the idea that two mutually replaceable populations of ISCs co-exist in the intestinal crypts[358]. Taken together, although several seemingly specific +4 cells markers were identified, gene expression profiling[232] and single molecule mRNA fluorescence *in situ* hybridization (FISH) analysis[143] revealed rather broad expression of these genes throughout the crypt and overlapping expression with Lgr5 in CBC cells. It is therefore difficult to unambiguously distinguish the two ISCs populations based on expression of their markers. Nevertheless, the hierarchy between these two populations within response to injury seems to be clear. In

a recent study, Zou and colleagues used rotavirus infection to specifically damage differentiated cells, leaving all ISCs intact. They showed that when Lgr5$^+$ CBC cells remain intact, they completely reconstitute the epithelium without the need of +4 quiescent ISCs activation[429].

Plasticity of TA progenitors and differentiated cell lineages

Until recently, it has been thought that the reserve pool of ISCs is specifically associated with the +4 position within crypt. However, following studies have shown that upon excessive damage of the intestinal epithelium, progenitors of secretory cell lineages may dedifferentiate to stem-like cells and recover the tissue (reviewed in[10]; Figure 5). Maintenance of an undifferentiated and proliferative state of secretory progenitors (TA cells) requires, *inter alia*, activity of the Notch signaling pathway[344]. Canonical Notch ligand Dll1 is expressed in a subset of TA progenitors which were thought to be a rapidly proliferating cell population that promptly differentiate to secretory lineages. However, sorted Dll1$^+$ cells are after exogenous Wnt stimulus able to give rise to organoids containing Lgr5$^+$ cells[381], which indicates at higher plasticity than was thought. Sex-determining region Y (SRY)-related high-mobility-group box 9 (Sox9) belongs to a family of Sox transcription factors that can modulate proliferation and differentiation of progenitor cells. In the small intestine, Sox9 is highly expressed in LRCs, TA progenitors, and enteroendocrine cells and less abundantly also in CBC cells[76]. Although under physiological conditions Sox9$^+$ cells are not able to form organoids *in vitro*, they exhibit increased ability to form organoids upon irradiation[382] or insulin-like growth factor 1 stimulation[383]. The importance of Sox9 in tissue regeneration is supported by observation that Sox9 depletion from the intestinal epithelium results in loss of LRCs and impaired regeneration upon irradiation[291]. Surprisingly, some post-mitotic cell populations are also able to convert back to stem cell-like status after tissue damage (reviewed in[10]). It has recently been shown that Paneth cells are capable of re-activating the cell cycle after irradiation, while at the same time acquiring stem cell properties and suppressing expression of Paneth cell-specific genes[294, 421]. Interestingly, Paneth cells isolated from irradiated mice gave rise to organoids *in vitro*[421]. Apparently, Paneth cells may switch from a non-dividing supportive cells to actively proliferating cells and contribute to intestinal tissue regeneration upon irradiation[294]. Recent studies have suggested that a subset of mature enteroendocrine cells and goblet precursors[145, 416] as well as precursors of absorptive enterocytes[367] are also able to regain ISCs identity when Lgr5$^+$ cells are ablated. To

conclude, all crypt cells exhibit a greater or lesser degree of plasticity with the ability to regenerate intestinal epithelium or replace the Lgr5^{+} cell pool upon injury.

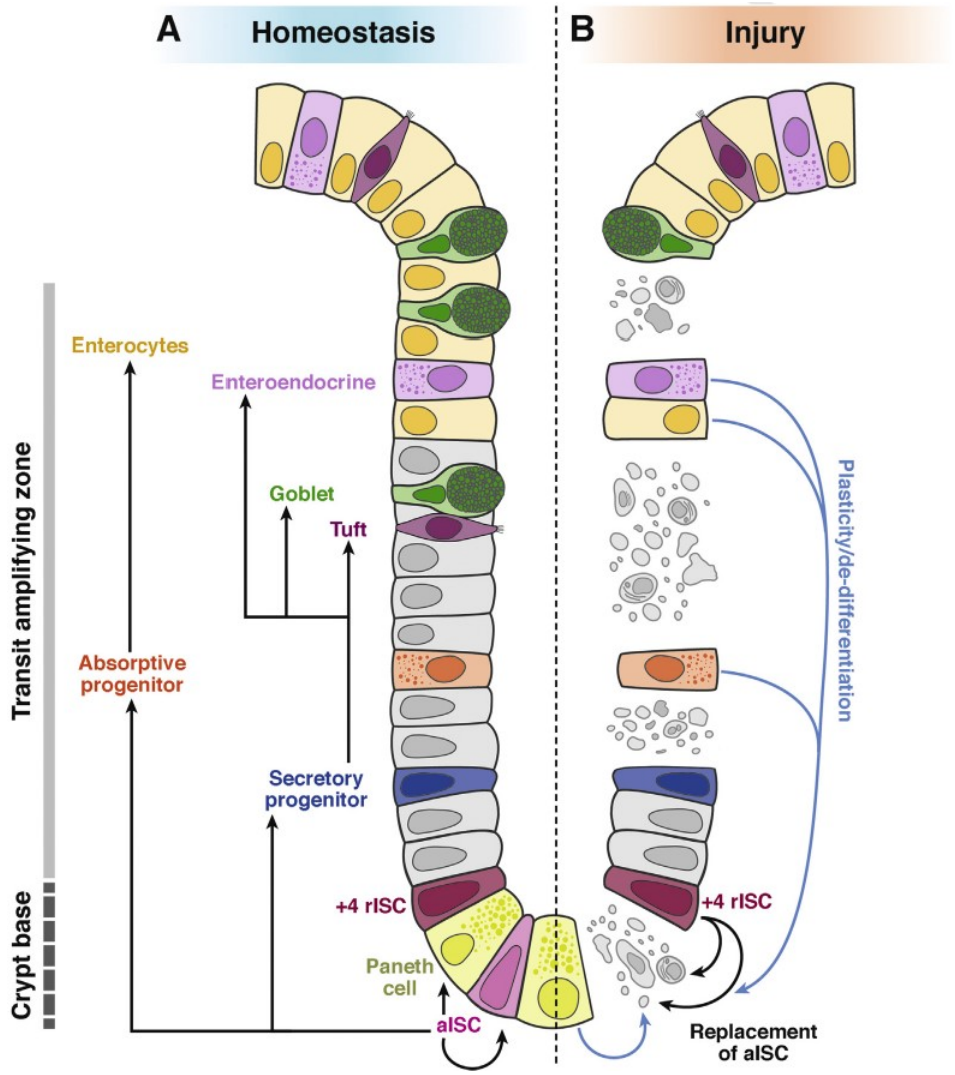


Figure 5 | Crypt cells in homeostasis and during regeneration upon injury.
(A) Actively dividing intestinal stem cells (aISCs) mediate renewal of the intestinal epithelium in homeostatic conditions. Rarely dividing reserve ISCs (rISCs) occupy the +4 position within the crypt and minimally contribute to the tissue maintenance. (B) Upon injury causing loss of aISCs, rISCs start to rapidly divide to refill the aISCs pool and regenerate the tissue. Progenitor cells in the TA compartment may also contribute to the regeneration, probably via dedifferentiation to aISCs (adopted from[10]).

1.2.3.2 Intestinal stem cell niche

The intestinal stem cell niche consists of numerous cell types of epithelial or mesenchymal origin and provides favorable microenvironment to self-renewing stem cells while ensuring essential signals for differentiation of progenitor cells. The intestinal crypts, originally named crypts of Lieberkühn after their discoverer Jonathan Nathanael

Lieberkühn, contain 14-16 Lgr5^{+} stem cells[339] interspersed among Paneth cells. Earlier models suggested asymmetrical division of Lgr5^{+} ISCs that resulted in one stem and one TA cell. However, recent findings support a model in which Lgr5^{+} ISCs divide symmetrically and lately adopt the TA fate or retain their stemness. The ISCs progeny compete for limited space within the niche and the adoption of TA fate probably depends on their contact with Paneth cells and position within the crypt. Current "neutral competition" model therefore suggest that ISCs are in a permanent competition when the number of Lgr5^{+} cells is limited by available surface of Paneth cells in the crypt[196, 339]. This model is supported by short- and long-term lineage tracing experiments in *Ah-Cre/Rosa26-Confetti* mice, a reporter mouse model that enables random activation of green, yellow, red, or blue fluorescent protein in most cells (including intestinal stem cells), but not in Paneth cells[138, 339]. Obtained data revealed that crypts tend to drift towards clonality (Figure 6), i.e. that each crypt becomes clonal within 3 months as a result of neutral competition between Lgr5^{+} ISCs[339]. Moreover, in a follow-up study Ritsma and co-workers showed that Lgr5^{+}

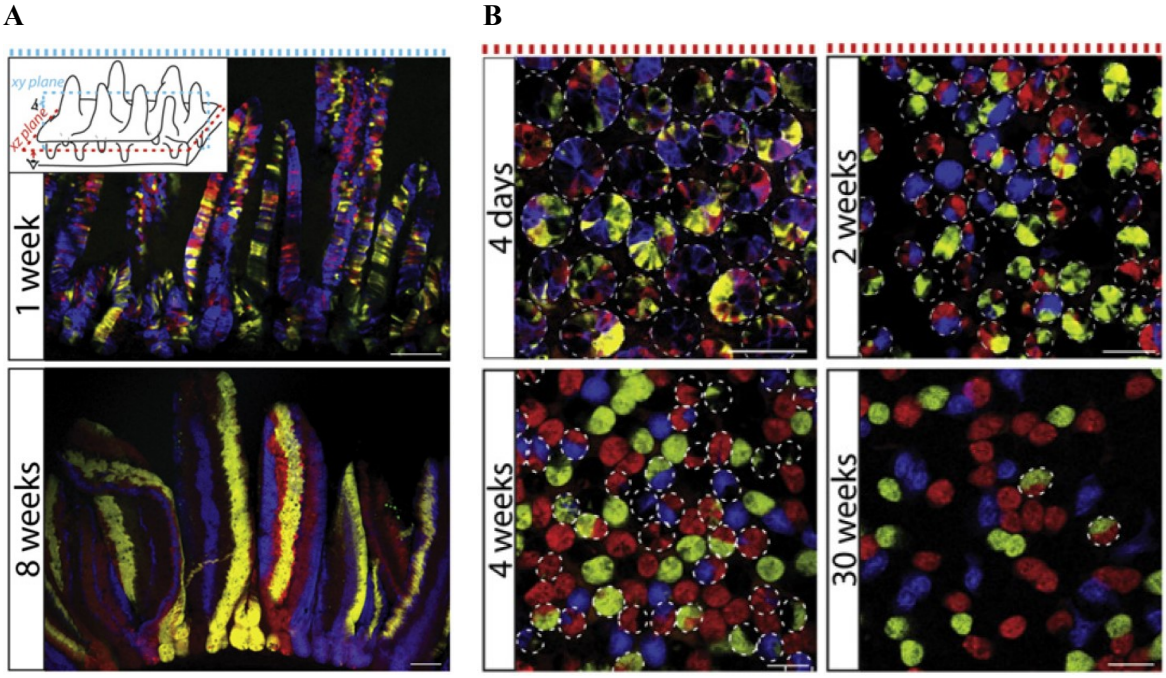


Figure 6 | Long-term lineage tracing reveals clonal character of the small intestinal crypts. Ah-Cre/Rosa26-Confetti mice enable inducible expression of GFP, YFP, RFP, or BFP in ISCs. Scheme of the small intestine indicates xy (blue) and xz (red) sectioning planes used for analysis. (A) The xy plane images were taken 1 and 8 weeks upon Cre activation, showing expansion of individual clones (yellow, red, and blue ribbons). (B) The xz plane images reveal the drift to clonality. At day 4 upon Cre activation, confocal cross-section of crypts shows heterogenous labeling; at later timepoints, crypts become more homogenous and, finally, labeled with only one color or completely unlabeled. Non-clonal crypts are indicated by white dashed circles; scale bar: 100 μm (adopted from[339]).

cells at the crypt bottom ("central cells") exhibit three times higher probability of colonizing the whole crypt than Lgr5$^+$ cells positioned at the border of the niche ("border cells")[287], further supporting the model of position-dependent fate adoption.

As mentioned earlier, the whole network of signaling pathways is responsible for maintaining ISCs and TA progenitors. Numerous cell types have been identified to contribute to the small intestinal niche by supplying the CBC cells with various growth factors and other signaling molecules. Although Wnt signaling is undoubtedly indispensable for intestinal stem cells maintenance[268], the dispensability (and to a certain extent also the origin) of individual sources of Wnt ligands is questionable. ISCs neighboring Paneth cells provide Wnt3, Notch ligand Dll4, EGF and transforming growth factor α (TGFα), which are all essential for ISCs culture *in vitro*[310]. Although Paneth cells were considered as indispensable for ISCs maintenance, their ablation from the epithelium had no effect on the crypt morphology *in vivo*[60, 164]. Alternative sources of Wnt ligands and niche-supporting growth factors have been identified in the mesenchyme[94]. Subepithelial myofibroblasts, which are closely associated with the crypt base, produce Wnt2b[65, 94] and other growth factors and cytokines (reviewed in[276]) to support ISCs proliferation. However, even simultaneous elimination of Wnt-ligand secretion in myofibroblasts and Paneth cells had no effect on stem cells proliferation or differentiation and tissue morphology, suggesting a higher level of redundancy of niche-maintaining factors production in the tissue[301]. Colonic stem cell niche is maintained by a Reg4$^+$ subpopulation of cKit/CD117$^+$ goblet cells, also termed Paneth-like cells[65]. These cells are, similarly to Paneth cells, adjacent to stem cells in the crypt base, produce EGF, Dll1, and Dll4 ligands, and facilitate organoid formation from Lgr5$^+$ cells *in vitro*[295]. However, production of Wnt proteins has not been observed in Reg4$^+$ cells[65]. The source of Wnt ligands in the colon is probably represented by glioma-associated oncogene homolog 1 (Gli1)-expressing mesenchymal cells, as the inhibiton of Wnt secretion from these cells results in loss of colonic stem cells and disruption of epithelial integrity[55]. Moreover, Gli1$^+$ cells seem to serve as a reserve population of Wnt-producing cells also in the small intestine, as they contribute to the pool of Wnt ligands when secretion from Paneth cells is blocked[55].

Crypt fission and ectopic crypt formation

Maintenance of the intestinal epithelium is driven not only by proliferation of progenitors in crypts, but by generation of new crypts as well[345]. The amount of crypts increases via crypt fission, a process by which a single crypt divides to two daughter crypts.

This is essential for intestinal elongation in postnatal development and maintenance of the epithelium in adulthood[43, 371]. The frequency of crypt fission is higher in young animals and becomes attenuated with age[131, 209, 345]; however, it is still active in adulthood during regeneration, e.g. upon partial resection or irradiation[36, 133, 371]. Importantly, crypt fission is restored in cancer to propel tumor growth[6, 132, 279]. The process of normal crypt fission has been described using small intestinal organoids. The site of fission initiation is apparently established in the stem cell niche between two Paneth cell-rich regions separated by a cluster of $Lgr5^+$ ISCs. Interestingly, the biomechanical properties of Paneth and stem cells play crucial role in the process – stiffer and more adhesive Paneth cells define the fission site, while softer and less adhesive $Lgr5^+$ cells enable shape changes and expansion of the crypt[185]. Although numerous studies describe the significance of crypt fission in tumor growth, molecular mechanisms responsible for the crypt fission reactivation remain unclear.

Aberrant crypt formation outside the "standard" crypt compartment accompanies morphological changes in the intestinal tissue after inactivation or transgenic expression of certain genes[102, 377] and is a defining histologic feature for some types of intestinal tumors (reviewed in[341, 370]). The so-called ectopic crypts are newly formed abnormally positioned crypts that lost their orientation to the underlying muscular layer of mucosa[370]. Several genes encoding regulatory proteins of several signaling pathways have been identified to underlie the ectopic crypt formation. In mice, transgenic expression of BMP inhibitor noggin throughout the intestinal epithelium led to formation of abnormal epithelial invaginations containing proliferating cells (Figure 7A, B) and expressing crypt-specific/Wnt target genes, e.g. *c-myc* and *EphB3*. These structures further developed to numerous ectopic crypt-villus units with perpendicular axis orientation towards the original crypt-villus axis. It was further proposed that inactivating mutations in components of the BMP signaling pathway influence sensitivity of epithelial cells to mesenchymal BMP signals. Consequently, loss of the inhibitory BMP signals leads to the formation of ectopic crypts that later progress to polyps and neoplasia[102]. Similar phenotypical changes were observed in transgenic mice expressing pan-hedgehog inhibitor hedgehog interacting protein in intestinal epithelial cells[134]. However, the ectopic crypt formation might be the result of perturbed BMP signaling which is controlled by Shh produced from mesenchymal cells[289]. Formation of ectopic crypts was also observed upon simultaneous activation of the Wnt and nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) signaling pathways. It was suggested that the aberrant "pouches" of proliferating cells originated

from non-stem epithelial cells which dedifferentiated to tumor-initiating cells[329]. Finally, transgenic expression of the Wnt target gene *Ascl2* in intestinal epithelial cells induced formation of ectopic crypts on villi accompanied by hyperproliferation of the crypt compartment (Figure 7C, D)[377].

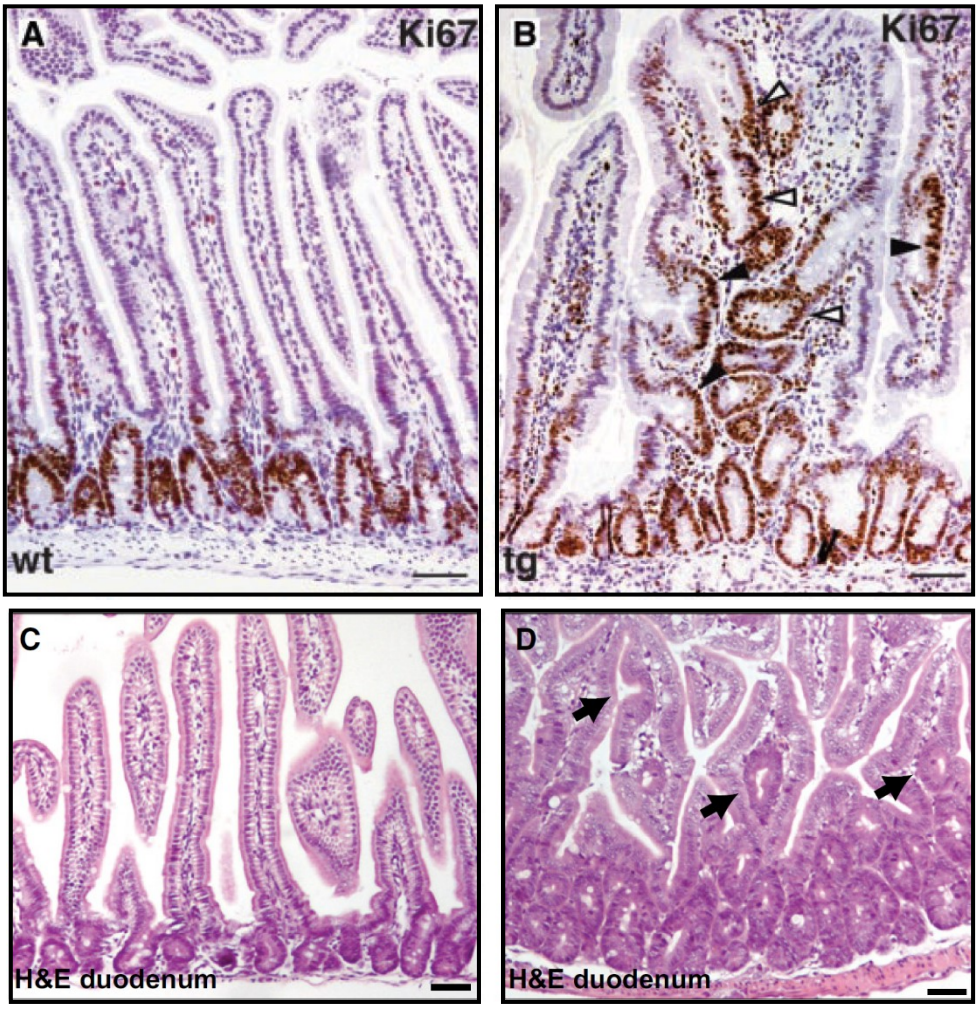


Figure 7 | Ectopic crypts formed in the small intestine.
(A) Histological staining of small intestinal sections from wild-type (wt) mouse. Proliferating cells stained with Ki-67 (brown nuclei) are found only in the crypt compartment. (B) Ectopic crypt-villus units are formed in *noggin* transgenic mice (tg). Ki-67-positive cells aberrantly appear outside the crypt compartment in crypt-like pockets. Black arrowheads indicate epithelial invaginations, white arrowheads indicate scattered crypts in stroma; scale bars = 0.1 mm (adopted from[102]). (C) Hematoxilin and eosin (H&E) staining of small intestinal sections from wt mouse show normal tissue morphology. (D) H&E staining of small intestinal sections from *Ascl2* transgenic mouse show a hyperplastic crypt compartment and branched villi. Black arrows indicate ectopic crypts; scale bars = 50 µm (adopted from[328]).

1.2.4 Colorectal cancer

Colorectal cancer (CRC) represents one of the most frequently diagnosed cancer types in developed countries with a death rate over 30 %. The lifetime risk of CRC is approximately 5 % and this number has been rising due to population ageing[331]. Based on the knowledge gained from numerous independent studies, it can be assumed that cancer is a stem cell disease (reviewed in[284, 360]). It is therefore not surprising that many pathways which regulate normal stem cells self-renewal are also associated with cancer development and progression (reviewed in[177, 354]).

Wnt signaling in colorectal cancer

The vast majority of colorectal tumors is associated with initial mutations in genes that regulate Wnt signaling. The most frequently mutated gene is the tumor suppressor *APC* (reviewed in[5, 167, 277]), a component of the β-catenin destruction complex. More than 60 % of *APC* mutations are located on exon 15 in the so-called mutation cluster region (MCR)[221]. In most cases, these mutations cause loss of C-terminal portion of the APC protein, which contains binding domains for AXIN2, β-catenin, and some interacting partners involved in cell polarity, microtubule assembly, and chromosome segregation (Figure 8)[220, 221]; the latter mentioned explains the link between *APC* mutations and chromosomal instability[72]. Mutations in other components of Wnt signaling might as well hyperactivate the pathway, although these are much less frequent than mutations in *APC*. *CTNNB1* gene (encoding the β-catenin protein) mutations have been found in approximately 5 % of CRCs. Short three-base deletion and point mutations in *CTNNB1* exons that encode functionally significant phosphorylation sites result in production of stabilized and therefore constitutively active β-catenin[136, 228]. Relatively rarely occur mutations in Wnt negative regulators *AXIN1* and *AXIN2*, or in the transcription factor *TCF4*[15, 193, 324, 362]. In addition to mutations or chromosomal rearrangements, epigenetic changes in components modulating Wnt signaling have been described (reviewed in[316]). Epigenetic silencing of *AXIN2* by excessive methylation of its promoter was observed specifically in colorectal tumors with microsatellite instability (MSI)[170]. Promoter hypermethylation was also found in several Wnt antagonists, such as DKK1[2], SFRP1/2/4[352], and WNT inhibitory factor 1[213]. Of note, several Wnt target genes have been linked to progression of colorectal tumors. Cyclin D1 and c-myc contribute to tumor growth and malignant progression[30, 107, 365]. Overexpression of CD44, a positive regulator of Wnt signaling[326], was observed already in earliest

hyperplastic lesions of colorectal tumorigenesis[404]. Matrix metalloproteinase 7 (Mmp7 or matrilysin) is overexpressed in up to 80 % of human CRCs[420] and was shown to have important function in adenoma growth, invasiveness, and metastasis[278, 406, 412]. To conclude, hyperactivation of the Wnt signaling pathway underlies initiation of intestinal tumors via stabilization of β-catenin and subsequent activation of the Wnt target genes, which propels intestinal epithelium transformation, tumor progression, and invasiveness.

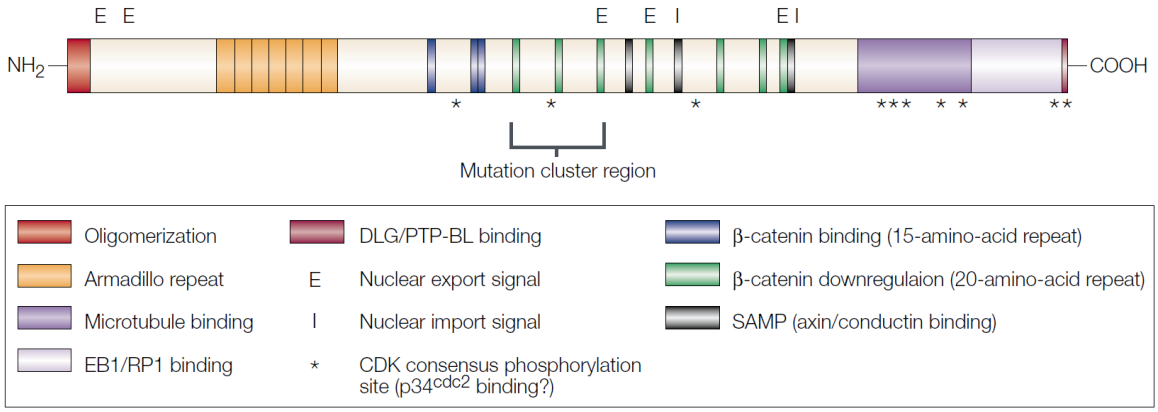


Figure 8 | The adenomatous polyposis coli (APC) protein. The scheme of APC protein shows a conserved region of Armadillo repeats and protein binding domains that interact with tubulin, microtubule-associated protein EB1, discs large (DLG), β-catenin, and axin/conductin. APC also contains five nuclear export signals (E) and two nuclear import signals (I). Consensus phosphorylation sites for cyclin-dependent kinase (CDK) are indicated by asterisks; the mutation cluster region, a portion of the *APC* gene sequence which is a subject of majority (over 60 %) of somatic mutations, is indicated by a black clamp (adopted from[74]).

Progression and histopathology of colorectal tumors

The onset of CRC might be sporadic or underlied by a hereditary cause. In both cases, the earliest stages of intestinal adenomas appear as the so-called aberrant crypt foci (ACF)[80], microscopic lesions that precede formation of epithelial neoplasia[248] which are generally associated with mutations in *APC*[149]. Following activating mutation in Kirsten rat sarcoma viral oncogene homolog (*KRAS*) enhances Wnt signaling and thus drives the adenoma growth[148, 337] (reviewed in[166]). Interestingly, lineage-tracing experiments in mice revealed dramatic acceleration of crypt fission upon *KRAS* mutation[340], suggesting that fission of transformed crypts is essential for colorectal adenoma growth[279, 409]. Subsequent sequential accumulation of additional mutations, including at least one oncogene and several tumor suppressor genes, or allelic losses promote polyp progression to malignant stage[388, 402]. The most commonly mutated genes include tumor protein 53 (*TP53*), *SMAD2*, *SMAD4*, phosphatase and tensin homolog (*PTEN*), and PI3K catalytic subunit α (*PIK3CA*)[74, 77, 290, 300, 388]. Concordantly, whole genome sequencing data of colorectal

cancer specimens revealed few commonly mutated genes and a large group of less frequently occurring amino acid-altering mutations[410].

Colorectal tumors develop through several well-defined histopathological stages including low- and high-grade dysplasia, adenoma, and invasive adenocarcinoma (Figure 9); individual stages are characterized by alterations in cell morphology[388] (Figure 10). Macroscopically, colorectal adenomas may be classified as elevated, flat or depressed, with elevated adenomas ranging from pedunculated polyps with a long stalk lined by normal mucosa to sessile polyps located on the surface of the mucous membrane. From the histopathological point of view, adenomas split into four groups: tubular, villous, tubulovillous, and serrated. Tubular adenomas are usually pedunculated and globular, but might also be flat. They are formed by irregularly arranged dysplastic tubular structures which account for at least 80 % of the luminal surface. Villous adenomas are often large and predominantly sessile and have a "velvety" surface. Microscopically, thin protrusions lined by dysplastic epithelium cover at least 80 % of the luminal surface. Tubulovillous adenomas display a mixture of tubular and villous morphology with the ratio between 80/20 and 20/80 %, respectively. Noteworthy, villous adenomas in human represent a more progressed stage along the path to fully developed CRC, which is reflected by deteriorated prognosis. Finally, serrated adenomas are hyperplastic polyps with luminal surface containing numerous prominent epithelial cells, which resemble a saw blade; this

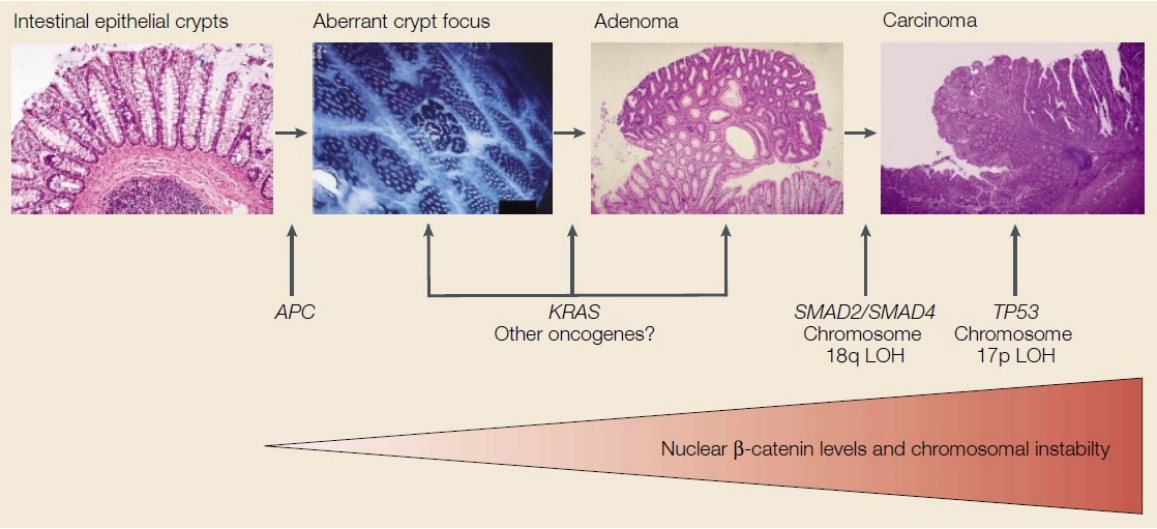


Figure 9 | Progression of colorectal tumors.
The earliest detectable stage of colorectal neoplasia is a microscopic lesion called aberrant crypt focus (ACF), which is linked to mutations in the *APC* gene. An additional mutation in *KRAS* promote formation of adenoma and following mutations in other genes, such as tumor suppressor p53 or *SMAD*, facilitate progression to malignant carcinoma (adopted from[74]).

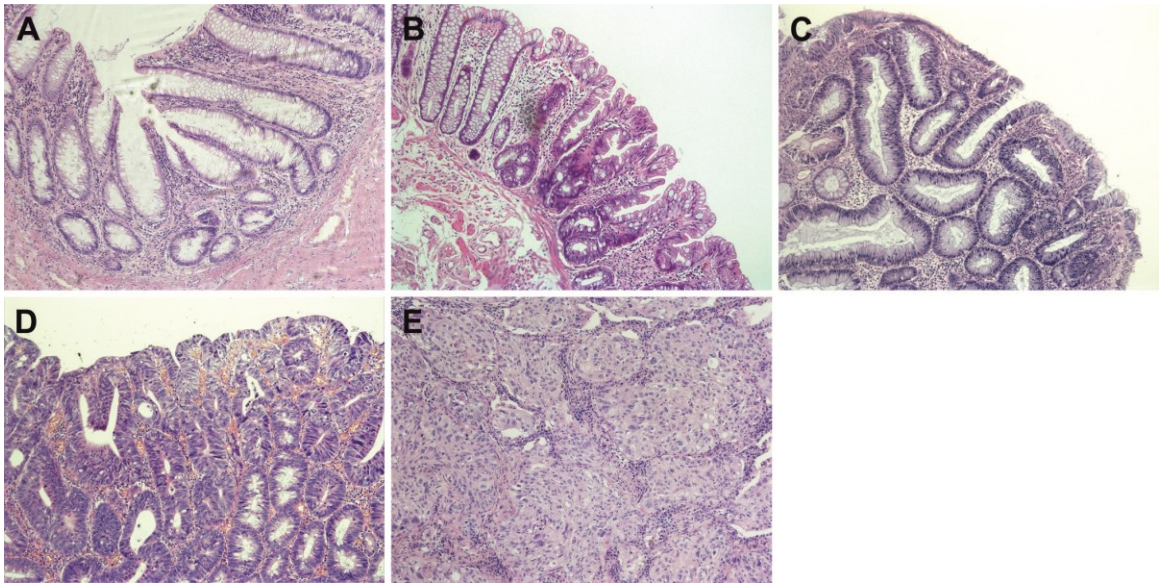


Figure 10 | Histopathology of colorectal tumors.
Paraffin sections of human colorectal neoplasia at different stages of progression from the healthy colon to carcinoma. (A) Control colonic epithelium with healthy crypts; (B) hyperplastic colonic epithelium contains prolonged crypts with serrated surface facing the lumen of colon; (C) cells in the low grade dysplasia have slightly enlarged nuclei oriented to basal membrane; (D) high grade dysplasia is characteristic by prominent nuclear stratification; (E) colorectal carcinoma (unpublished pictures provided by J. Svec).

morphological feature is primarily due to defective apoptosis[90]. Based on histological morphology, serrated adenomas are according to World Health Organization classified into three types: hyperplastic polyps (HP), sessile serrated adenomas/polyps (SSA/P), and traditional serrated adenomas (TSA; reviewed in[334]). TSA are minor subtypes of colorectal carcinomas, however, they are precursors of biologically aggressive types of colorectal tumors. Of note, a typical feature of TSA is the presence of ectopic crypts (reviewed in[22]).

1.2.4.1 Hereditary CRC syndromes

Although approximately one third of patients diagnosed with CRC have a family history of cancer, suggesting the presence of a hereditary cancer-causing component, only 5-10 % of CRC cases are linked to particular gene mutations. The hereditary CRC syndromes are divided into two main groups, polyposis and nonpolyposis syndromes (reviewed in[122]). The group of hereditary nonpolyposis colorectal cancer (HNPCC) syndromes includes Lynch syndrome, Lynch-like syndrome, and familial colorectal cancer type X (FCCX). Lynch and Lynch-like syndromes account for about 3-5 % of all CRCs, exhibit MSI and carry mutations in genes modulating colonic stem cells proliferation and mismatch repair (MMR; reviewed in[122, 255]). The patients are also at higher risk of

developing tumors in other tissues, mainly in ovaries, endometrium, and stomach (reviewed in[122]). The second group of the hereditary CRC syndromes is represented by the familial adenomatous polyposis (FAP; reviewed in[73]) syndrome which accounts for less than 1 % of all CRCs . Individuals affected by the FAP syndrome carry a germline mutation in one *APC* allele which causes expression of truncated APC protein. Owing to the high frequency of spontaneous mutation in the second "healthy" allele, FAP patients develop hundreds to thousands benign colorectal polyps in the colon and rectum by the age of twenty years. As a result of successive accumulation of mutations in *KRAS*, *TP53*, or other tumor-promoting genes[74, 165, 240], the polyps inevitably progress to carcinoma by the age of 35 to 40 years (reviewed in[83]).

1.2.4.2 Mouse models of CRC

For the purpose of studying mechanisms leading to initiation and progression of colorectal tumors, genetically modified mouse strains that resemble mutations found in patients were generated (reviewed in[359]). An exhaustive description of all genes that have been modified in mice to study CRC would be beyond the scope of this thesis, therefore only mouse models and experimental approaches relevant to the thesis are mentioned. The most common mouse models used for studying intestinal cancer biology have been described in detail in a recent review[266].

Mouse Apc models are greatly suitable to study CRC, as they share 90% similarity at the protein level and all motifs, which have been characterized in human APC, are well conserved in mice[349]. Moreover, the gene signature in tumors from mice with mutant *Apc* and in human tumors with germline *APC* mutations are similar[86]. Multiple intestinal neoplasia (Min) mice ($Apc^{+/Min}$) is a frequently used mouse strain that carries a germline nonsense mutation in one allele of the *Apc* gene and therefore expresses a truncated, 850 amino acids long Apc protein[230, 349]. Similarly to FAP patients, the second *Apc* allele becomes randomly mutated and consequently multiple intestinal polyps are developed in adulthood[249] (Figure 11). Homozygosity for this mutation is lethal at early stages of embryonic development. Of note, polyps in $Apc^{+/Min}$ mice are predominantly located in the small intestine whereas polyps in FAP patients occur mainly in the colon; other tissues are affected only rarely in mice. Many other mouse strains carrying *Apc* mutations were generated which carry various *Apc* germline mutations[424]. It was proposed that the level of proliferation, differentiation, and apoptosis in adult stem cells depends on tissue-specific levels of β-catenin; consequently, the susceptibility to tumorigenesis and distribution of

tumors depends on particular *Apc* mutations and varies in different Apc-deficient mouse strains[85, 424].

An alternative tool are mouse strains harboring conditional alleles allowing temporally regulated tissue-specific gene inactivation when crossed to a Cre recombinase-expressing strain. Cre is a DNA recombinase of virus origin which specifically recognizes a 34 bp palindromic DNA sequences termed the *loxP* sites[313]; based on the *loxP* sites orientation, Cre recombines or inverts the DNA sequence between two sites. Cre activity may be spatially determined using various endogenous promoters that drive its expression in particular cell types. Improved version of the enzyme was prepared by Cre fusion with the mutated form of human estrogen receptor (Cre-ERT2)[68]. Cre-ERT2 is retained in the cytoplasm and therefore not active; however, upon binding of 4-hydroxytamoxifen (4-OHT, a metabolite of tamoxifen), Cre-ERT2 translocates to the cell nucleus and can induce desired DNA rearrangements (reviewed in[236]). A large number of mouse strains carrying modified *Apc* gene alleles have been generated which allow for various shortening of the Apc protein. In this theses, mice harboring conditional knock-out (cKO) alleles of the *Apc* gene with floxed exon 14 ($Apc^{cKO/cKO}$) were used[182]. By crossing these mice to strain expressing Cre recombinase under control of murine *villin1* promoter (*Villin-CreERT2*)[62], Apc inactivation is induced in the entire intestinal epithelium, resulting in crypt hyperplasia observable within several days[3]. These mice, however, die within 3-5 days due to extremely extensive damage of the epithelium. Alternatively, $Apc^{cKO/cKO}$ mice can be crossed with the *Lgr5-EGFP-IRES-CreERT2* strain, which drives Cre expression from the *Lgr5* gene promoter. In this strain, Apc is inactivated specifically in intestinal stem cells and expansion of the crypt compartment is slower; within 1-2 weeks, multiple microadenomas are formed which progress to macroscopic adenomas in 3-5 weeks[13]. Finally, a model of sporadic CRC can be induced in $Apc^{cKO/cKO}$ mice by intrarectal injection of adenoviral particles encoding the Cre recombinase, which develop isolated tumors in the distal colon within cca 18 weeks after infection (reviewed in[120]).

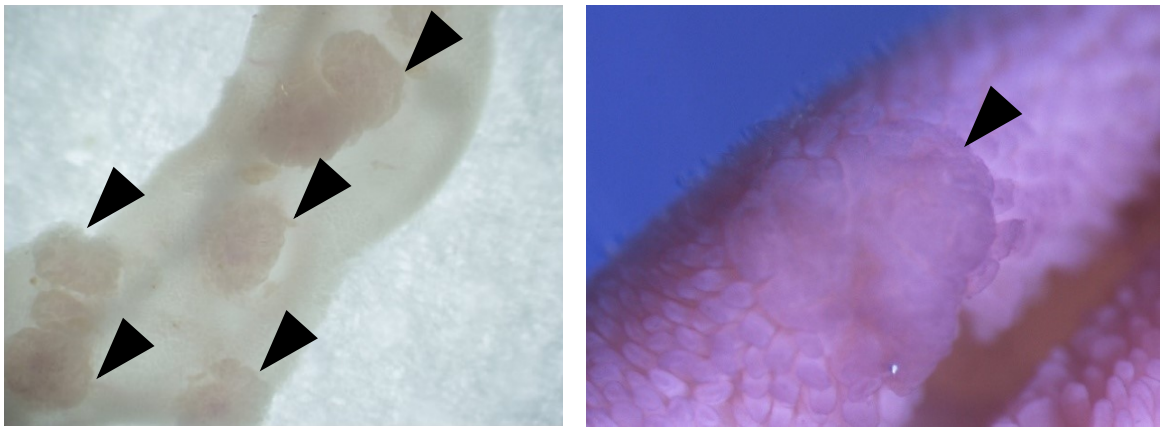


Figure 11 | Intestinal tumors in *Apc*$^{+/Min}$ mice.
Apc$^{+/Min}$ mice develop numerous tumors in the small intestine and sometimes (usually one large tumor) in the colon. Left, a macroscopic picture of the mouse intestines with tumors (indicated by black arrowheads); right, stereomicroscopic image of the inner intestinal surface with a protruding polyp (indicated by a black arrowhead; unpublished pictures provided by B. Fafilek).

1.2.4.3 Colitis-associated colorectal cancer

Inflammatory bowel disease (IBD) is a relapsing chronic inflammatory condition of the intestinal tract, which belongs to the group of autoimmune diseases. Two most common subtypes of IBD are Crohn's disease (CD) and ulcerative colitis (UC; reviewed in[378]). CD typically affects the entire digestive tract, with lesions located in the small intestine and the proximal part of the colon (reviewed in[70]). Inflammation in UC is primarily localized to the rectum and continuously expands in the proximal direction through colon, but only occasionally reaches the small intestine (reviewed in[373]). Colitis-associated colorectal cancer (CA-CRC) accounts for 10-15 % of IBD patients deaths; however, it causes only 1-2 % of all CRCs (reviewed in[211]). In CA-CRC, tumors are located in the areas of colon with active inflammation and develop, as in "standard" CRC, from dysplasia to carcinoma due to accumulation of mutations in epithelial cells (reviewed in[378]). Interestingly, the initial mutation in CA-CRC does not alter the *APC* gene, but it inactivates the tumor suppressor *TP53*[211] (Figure 12); however, early activation of Wnt signaling has been described as critical to the process of colitis-to-cancer transition[321]. Of note, recent studies reported different development of CA-CRC in patients exhibiting similar patterns of the inflammation, which suggests that other factors contribute to the CA-CRC

progression, possibly polymorphisms or mutations in low penetrance disease susceptibility genes[378].

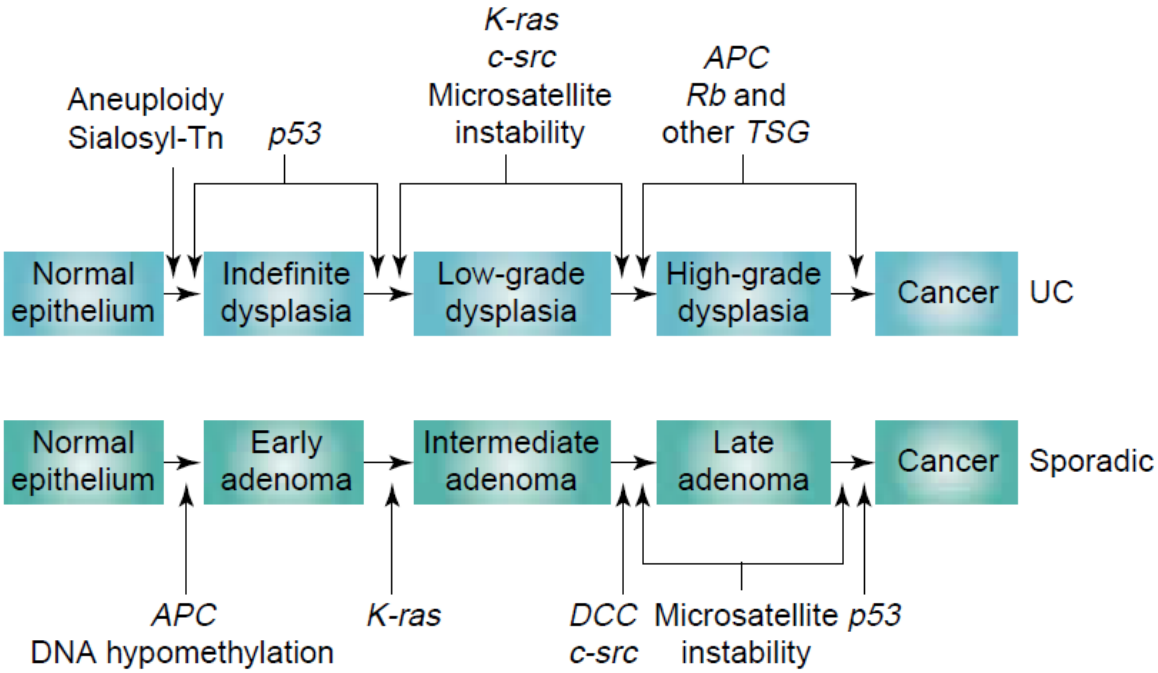


Figure 12 | The sequence of obtained mutations differs between sporadic colorectal cancer (CRC) and colitis-associated cancer (CAC).
The scheme shows comparison of colitis-associated and sporadic colorectal cancer. Both types of cancer develop through accumulation of multiple mutations; nevertheless, the sequence of mutation in *APC* and *TP53* is different. DCC, deleted in colorectal cancer; Rb, retinoblastoma; TSG, tumor suppressor genes; UC, ulcerative colitis (adopted from[285]).

Mouse models of CA-CRC

As the genetic approaches are time consuming, expensive, and affect the colon to a much lesser extent than the small intestine, methods for CRC induction based on chemical treatments of experimental mice have been invented. In the azoxymethane (AOM)/dextran sodium sulfate (DSS) model (AOM/DSS), DNA damage is followed by colitis leading to generation of colonic tumors[361]. AOM is a procarcinogen which upon metabolic activation induces formation of O6-methyl-guanine which results in frequent point mutations during DNA replication. Subsequent administration of DSS causes inflammation, which enhances the tumor incidence and significantly accelerates the ACF-adenoma-carcinoma pathway (reviewed in[256]). Furthermore, AOM/DSS-induced tumors exhibit similar histopathological aspects to human CA-CRC, such as localization in the distal colon and invasivity[54].

1.3 Transcription factors linked to the Wnt signaling pathway

As already mentioned above, aberrant activation of the Wnt signaling pathway underlies initiation of the majority of the colorectal tumors. Additionally, the Wnt/β-catenin regulated genes in the intestinal epithelium encode proteins that perform a wide range of functions important for the intestinal homeostasis. This thesis deals with two transcription factors which are linked to the Wnt signaling pathway, msh homeobox 1 (MSX1) and hypermethylated in cancer 1 (HIC1). While studies of the MSX1 protein function in human cancer brought often contradictory conclusions, HIC1 has been described by numerous independent research groups as a potent tumor suppressor. This chapter describes main properties and functions of these two transcription factors in embryonic development, intestinal homeostasis, and cancer.

1.3.1 Msx1 transcription factor

Msh homeobox 1 (*Msx1*) was described as a new member of mouse homeo box-containing gene family in 1989, at that time named *Hox7*, as it displayed striking similarity to the *Drosophila* Msh homeobox[111]. Msx1 has been studied extensively in mouse embryogenesis and human diseases associated with defective tooth development. Alterations in *MSX1* expression have been described in many types of human tumors; however, its function in intestinal tissue homeostasis or CRC remains unclear.

1.3.1.1 The *Msx1* gene

Msx1 is a member of the muscle segment homeobox (msh) gene family (reviewed in[21]) that belongs to the most conserved homeobox transcription factors in animals (reviewed in[112]). The *Msx1* gene is located on proximal end of chromosome 5 in mice[111]. In human, *MSX1* maps to 4p16.1 locus (on the short arm of chromosome 4) that is deleted in the Wolf-Hirschhorn syndrome[144], a human disease associated, *inter alia*, with craniofacial malformations. Both human and mouse *Msx1* genes consist of two exons interrupted by cca 1.6 kb long intron and share a very high degree of identity on both the DNA and protein levels[110].

The TATA-less *Msx1* promoter contains three regulatory regions responsible for spatial control of the complex *Msxl* expression pattern during embryonic development. The distal element localized 4 kb upstream from mouse transcriptional start (+1) drives *Msx1* expression within the first and second pharyngeal arches, which gives rise to (some)

craniofacial muscles, bones, and nerves, such as maxila and mandible (upper and lower jaw), chewing and pharyngeal muscles, facial and trigeminal nerves, nose, and middle ear bones. The proximal element localized 2.2 kb upstream from mouse transcriptional start drives *Msx1* expression in the third pharyngeal arch, dorsal neural tube, dermomyotome, and limb bud mesenchyme[201]. The last element is directly adjacent to mouse transcriptional start (+165/-106 bp) and supports *Msx1* expression in craniofacial and skull bones and nose; it has been proposed as the minimal *Msx1* promoter in mice[357].

Msx1 expression is further regulated by numerous transcription factors. The first attempt to characterize putative transcription factors binding sites was based on structural and functional analysis of almost 5 kb long sequence upstream from the translation start site. Computer analysis revealed four major regions including binding sites for some ubiquitous transcription factors (AP2, NF-κB), developmentally regulated transcription factors (MyoD, engrailed, bicoid), transcription factors involved in cell proliferation (c-myc, JunB), and also sites for autoregulation[91]. More advanced methods later confirmed, that these and many other transcription factors, such as Tcf4[218], FGF4[161], BMPs[81], Pax9[244], or SMAD8[24] induce or regulate *Msx1* expression, which will be discussed in following chapters.

1.3.1.2 Msx1 protein

The human and mouse *Msx* family consists of *Msx1*, *Msx2*, and *Msx3* genes (reviewed in[48]). The genes encode homeodomain-containing DNA-binding proteins that act as transcriptional regulators modulating morphogenesis during embryonic development. Msx1 is a 40-kDa protein of 297 amino acids (Figure 13) including 60 amino acids long homeodomain, which is 100% identical between mouse and human; overall, murine and human Msx1 proteins share 80% identity[110]. On the molecular level, Msx1 may act as a transcription activator or repressor, depending on cellular context and interacting partners.

Transcriptional regulation by Msx1 can be accomplished in several ways. Msx1 is a potent transcriptional repressor which can regulate transcription from both TATA-containing and TATA-less promoters via interaction with the core transcription complex[41]. The Msx1 homeodomain seems to be more important for interaction with other proteins than for binding to DNA and, interestingly, is dispensable for some Msx1 functions[41, 258]. Studies on myoblast differentiation revealed that MSX1 can regulate the gene expression also at the epigenetic level. In myoblast cells and developing limbs, MSX1 recruits a

polycomb repressive complex 2 (PRC2) to the nuclear periphery, which results in redistribution of H3K27me3 repressive mark; this proces is essential for cell differentiation[393, 394]. MSX1-mediated histone modification may be as well accomplished by association with G9a histone methyltransferase, which catalyzes synthesis of H3K9me2 repressive mark[395].

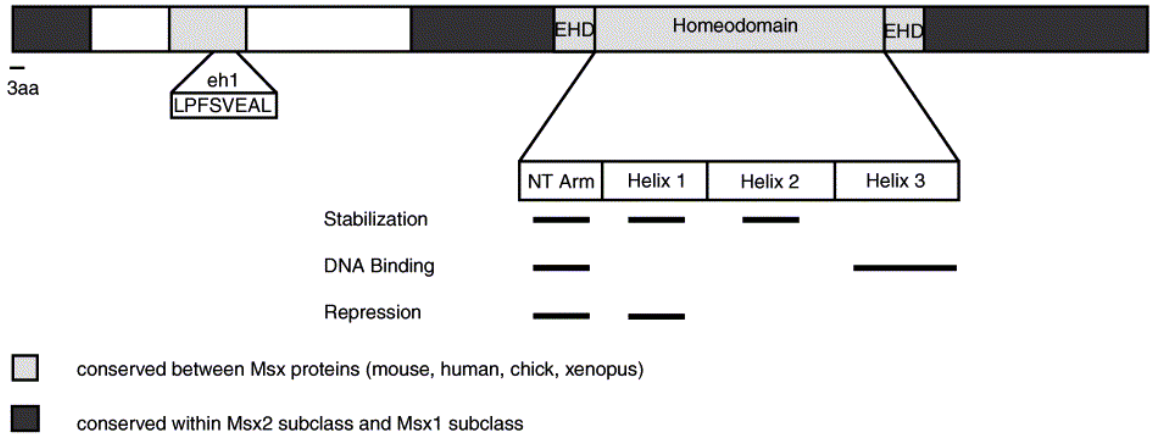


Figure 13 | Functional domains of Msx proteins.
The diagram shows domain composition of Msx proteins. Light grey boxes indicate conserved regions between Msx proteins in mouse, human, chick, and xenopus; dark grey boxes indicate conserved regions between Msx1 and Msx2 subclasses. Various functions of distinct regions within the homeodomain are depicted by black lines under the diagram (adopted from[212]).

1.3.1.3 Msx1 in embryonic development

Vertebrate Msx proteins are broadly expressed during embryonic development at diverse sites of epithelial-mesenchymal interactions and have important functions in organogenesis. Whereas *Msx3* expression is restricted to the dorsal neural tube[323, 399], *Msx1* and *Msx2* exhibit extensive and often overlapping expression in numerous embryonic tissues, such as limb buds, teeth buds, craniofacial structures, and heart[111, 127, 128, 184, 288], suggesting their functional redundancy. On the other hand, in some embryonic tissues *Msx1* and *Msx2* expresssion patterns are complementary, which points to different context-dependent functions[199].

The expresion pattern of *Msx1* in embryogenesis was first described by *in situ* hybridization staining of mouse embryos at different developmental stages. *Msx1* mRNA was detected already 6.5 days post coitum in extraembryonic tissue (amnion and ectoplacental cone) and in eight days old mouse embryos also in the rostral portion of neural fold-underlying mesenchyme and in the area of neural folds which later gives rise to neural crest cells[288]. At the embryonic day 9.5 (E9.5), *Msx1* was expressed in the neural crest,

neural tube, developing forelimb bud, visceral arches, and heart. At later stages, the *Msx1*-specific probe labeled areas of the central nervous system, interdigital mesenchyme in fore- and hindlimb buds, facial regions derived from neural crest, and dental papilla[111, 288]. Of note, *Msx1* is broadly expressed also postnatally, for example in the uterus, cerebellum, dorsal skin, lungs, ovaries, and testes. Interestingly, *Msx2* expression pattern is almost identical, however, the relative abundance varies in different tissues[199]. In adults, *Msx1* is typically expressed in progenitor cells where it represses transcription of pro-differentiation genes[69, 235, 343, 389]; upon differentiation, *Msx1* expression usually decreases[116, 393].

Further studies of the Msx1 function were performed in mutant or transgenic mouse strains. Mice homozygous for germline mutation in the *Msx1* gene manifested respiration deficiency and numerous developmental defects, such as the cleft palate, tooth agenesis, and abnormal morphology of craniofacial bones; these animals die within 24 hours after birth[311]. Interesingly, Msx2-deficient mice are viable, although they display defects in the skin, teeth, skull, and mammary gland development and also impaired chondrogenesis and osteogenesis[140, 312]. Mice harboring mutations in both *Msx* genes exhibited enhanced phenotype of the single mutants, e.g. anomalies in development of limbs, ventral body wall, craniofacial and cardiac structures, neural crest, and central nervous system[9, 100, 141, 184, 245] and thus die prenatally between the embryonic day 17 and 18 (Figure 14).

1.3.1.4 Msx1 in human disease

MSX1 deficiency in human causes pleiotropic phenotypes associated with non-syndromic tooth agenesis, non-syndromic cleft lip with or without cleft palate, Witkop syndrome (a disorder characterized by thin nails and absence of several teeth), and Wolf-Hirschhorn syndrome[151, 154, 239, 386], with the individual phenotypes associated with specific mutations. In-frame mutations affecting the MSX1 homeodomain predominantly cause tooth agenesis, which may be accompanied by other developmental defects, while mutations not affecting the homeodomain cause mainly nonsyndromic orofacial cleft. Truncating mutations affecting the homeodomain result in more severe phenotypes and are associated with previously mentioned syndromes[189]. Notably, in many studies the association between tooth agenesis and cancer has been described in patients carrying *MSX1* mutation[26]. This is, however, not so surprising, as regulatory proteins involved in embryogenesis are often reactivated in tumors (reviewed in[37]).

MSX1 properties have been studied also in human cancer; however, various studies often show contradictory views on the MSX1 functions. A few studies indicate that *MSX1*

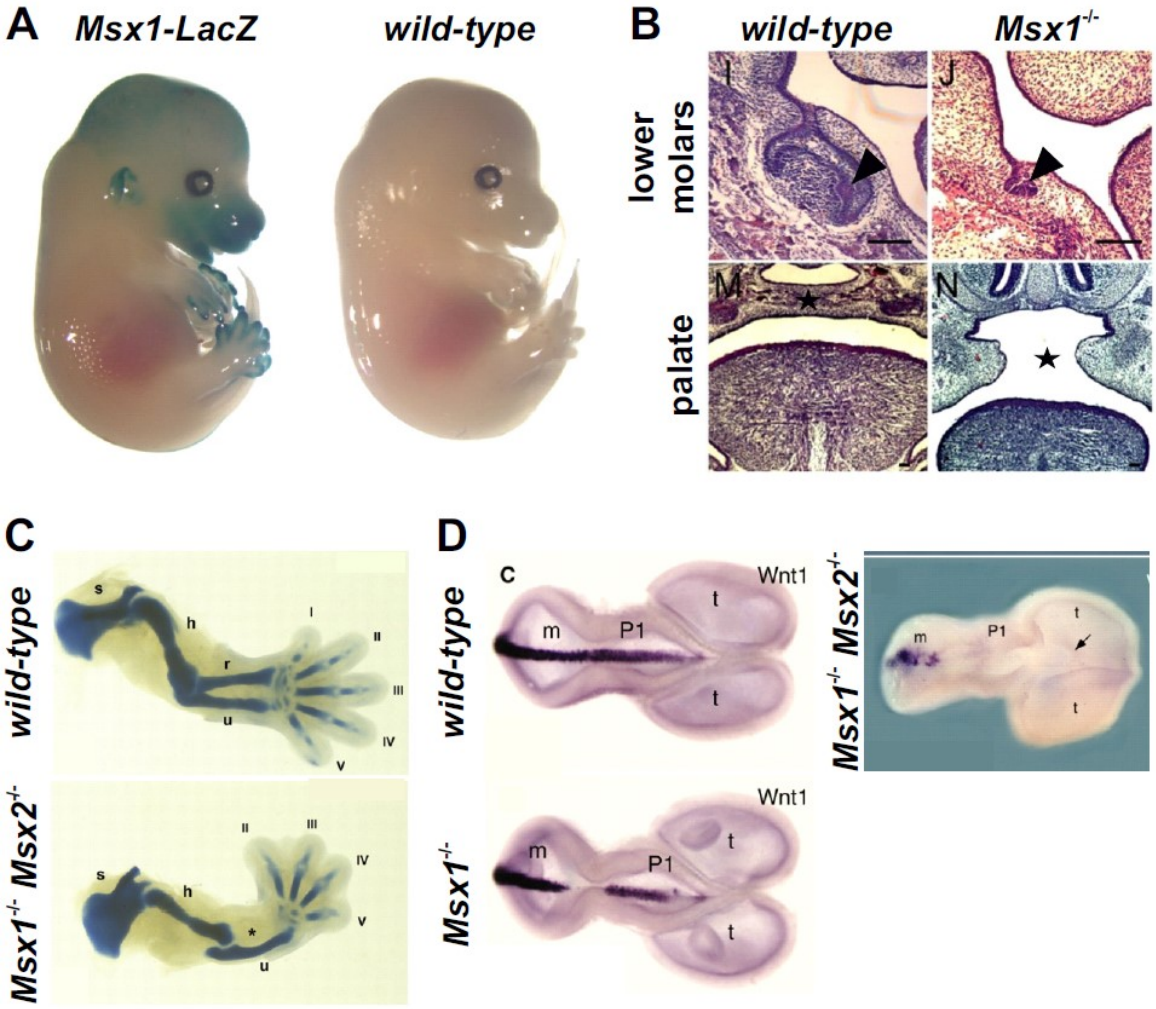


Figure 14 | *Msx1* expression in embryonic development.
(A) *Msx1* expression visualized by whole-mount X-gal staining of *Msx1-LacZ* mouse embryo. At embryonic day 14.5 (E14.5), *Msx1* is expressed in the limb buds, brain, ears, and craniofacial region (unpublished pictures provided by L. Janeckova). (B) Coronar sections of Msx1-deficient mouse embryo (E15.5) shows arrested molars at the stage of tooth buds and incomplete fuse of palate (adopted from[298]). (C) Skeletal preparations stained by Alcian blue show defective development of limb bones in Msx1-/Msx2-double-deficient embryos at E14.5 (adopted from[184]). (D) Whole-mount *in situ* hybridization of *Wnt1* mRNA reveals decrease of *Wnt1* expression in the midline region in Msx1-deficient mouse embryos and almost complete loss in Msx1-/Msx2-double-deficient embryos (adopted from[9]).

overexpression induces transformation[222, 234], whereas other studies describe MSX1 as a tumor suppressor[257, 322, 422]. Observed discrepancies possibly indicate tissue-specific and context-dependent functions of MSX1. For example, *MSX1* overexpression in human ovarian cancer cell lines leads to inhibition of cell proliferation and increased apoptosis[257]; moreover, the levels of *MSX1* expression seem to correlate with platinum drug sensitivity[26, 27]. *MSX1* is often silenced in cervical cancer cells, probably due to promoter hypermethylation, and its overexpression results in restoration of apoptosis, cell cycle

arrest, and inhibition of migration[258, 423]. On the other hand, *MSX1* expression was increased in breast cancer cells where it enhanced their invasive capabilities[61, 216], and also in pituitary adenomas[222] and liposarcomas[51].

Finally, several studies brought findings about *MSX1* expression in intestinal tumors. *MSX1* overexpression was detected in human CRC on the RNA and protein level and immunohistochemical staining revealed MSX1 signal in CRC cells with equal distribution between tumor center and invasive front[115]. Gene expression analysis combined with DNA methylation profile of colon adenocarcinoma and control tisssues revealed that *MSX1* expression negatively correlated with promoter methylation, i.e. that *MSX1* was hypermethylated and downregulated in colon adenocarcinoma[396]. These results bring opposite observation to studies that described *MSX1* overexpression in cancer cells[222, 234]; however, the authors suggest that the promoter hypermathylation may be a tumor-acquired feature. This would be in accordance with previous findings that *MSX1* overexpression represses proliferation of cancer cells[257, 422]. Of note, increased DNA methylation of the *Msx1* locus was described in mouse colonic epithelial cells that were exposed to colitis by AOM/DSS treatment[105]. However, there is otherwise a significant gap of knowledge regarding the Msx1 function in mouse intestinal tumors. Finally, the MSX1 function in cancer may be associated with metastasis since MSX1 seemingly facilitates the epithelial mesenchymal transition (EMT)[128, 299, 303], which is a prerequisite for migration of cancer cells from the primary tumor[118].

1.3.2 Hic1 transcription factor

Studies of the *HIC1* gene dates back to 1995 when Wales and colleagues sought genes located on the short arm of chromosome 17 (17p13.3), a region telomeric to the gene encoding the p53 tumor suppressor. This locus was, according to allelic loss data, suggested to harbor other tumor suppressor gene(s). In addition to the *HIC1* encoding sequence, Wales and colleagues also identified a p53-binding site in the *HIC1* regulatory region[391] and following functional studies confirmed direct p53-induced activation of *HIC1* expression[97, 391]. Moreover, it was later discovered that HIC1 is involved in a regulatory loop activated by DNA damage which results in p53 downregulation[56] and studies in mice revealed a functional synergism of Hic1 and p53 in tumor growth suppression[125]. Additionally, loss of HIC1 has been linked to the Miller-Dieker syndrome (MDS), a disease caused by chromosomal microdeletion of the 17p13.3 chromosomal region. Patients with

MDS suffer from hypotonia and lissencephaly (no gyrification of brain) and have extensive craniofacial defects[419]. Epigenetical silencing of *HIC1* by hypermethylation of its promoter has been observed in many human cancers, including colon, brain, lung, prostate, breast, and liver[71]. Nevertheless, *HIC1* expression exhibits a wide variety among cancer tissues and in some cases it is similar to matching healthy tissue. Of note, high production of HIC1 is characteristic for a specific type of intestinal tumors sensitive to chemotherapy[147].

In the mouse, *Hic1* is expressed in all adult tissues and exhibits an interesting expression pattern in developing embryos. *In situ* hybridization of *Hic1* mRNA revealed expression in many precursor areas that later give rise to tissues that are affected in patients with MDS, e.g. the first pharyngeal arch which becomes mandible, maxilla, and palate. *Hic1* is also expressed in the mesenchyme adjacent to budding epithelia of inner organs, such as nose, salivary glands, gut, and urogenital ridge. Furthermore, *Hic1* expression was detected in the lateral body wall, precartilaginous condensations, limb buds and developing kidneys[95]. *Hic1*$^{-/-}$ embryos are smaller than wild-type embryos, have severe developmental defects including acrania, exencephaly, cleft palate, limb abnormalities, and omphalocele and die perinatally[39]. Heterozygous *Hic1*$^{+/-}$ mice exhibit no defects during embryonic development and are viable; however, they develop numerous malignant tumors in many tissues due to epigenetical silencing of the "healthy" allele[126].

HIC1 is an evolutionarily conserved transcription repressor, which controls expression of several genes that are involved in proliferation, cell cycle regulation, development, tumor growth, and metastatic invasion[150, 292]. The HIC1-mediated transcriptional repression may be accomplished in several ways. The HIC1 N-terminal *broad-complex, tramtrack, and bric-à-brac* [BTB, alternative name *poxvirus and zinc finger* (POZ)] domain mediates oligomerization of the HIC1 polypeptide[374] and is able to autonomously repress transcription of the target genes without co-operation with histone deacetylases (HDACs; reviewed in[159]). The central part of the HIC1 protein harbors additional autonomously-active transcriptional repression domain with evolutionarily conserved protein binding motifs responsible for interaction with other proteins, e.g. with the general transcriptional co-repressor C-terminal-binding protein 1 (CtBP)[57]. Moreover, HIC1 in complex with its co-repressors binds to the regulatory region of the gene encoding HDAC Sirtuin 1 (*Sirt1*)[385]. Furthermore, HIC1 interacts with human polycomb-like proteins and recruits the PRC2 to the target genes[29]. Interestingly, HIC1 is able to indirectly inactivate transcription of its target genes that do not have HIC1-binding sites in their regulatory regions. This transcriptional repression was found in Wnt signaling, where

Valenta and co-workers identified interaction between HIC1 and β-catenin/TCF4 complexes. Upon binding to HIC1, the proteins were sequestered into nuclear structures called "HIC1 bodies", resulting in decreased Wnt signaling activity[374].

Hic1 role in the intestinal tissue has not been clearly defined, as only a few studies have dealt with Hic1 in the intestine. Some Hic1-regulated genes with functions in the mouse intestines have been described. Hic1 negatively regulates expression of *Atoh1* (alternative name *Math1*), a gene encoding a transcription factor which promotes secretory lineage specification during differentiation of progenitor cells in intestinal crypts (reviewed in[330]). Loss of *Atoh1* results in depletion of Paneth, goblet, and enteroendocrine cells[418] and increased tumor formation in $Apc^{+/-}$ mice[263]. Similarly, in human, ATOH1 antagonizes tumor growth by regulating proliferation and apoptosis, and in APC-deficient CRC tumors, *ATOH1* promoter is silenced by DNA methylation, which allows tumor cells to escape the cell cycle regulation and apoptosis[28, 231]. On the molecular level, Atoh1 upregulates expression of genes encoding $p57^{kip2}$ and $p27^{kip1}$ cell cycle inhibitors, thus arresting the cell cycle progression[163]. Although the direct effect of Hic1 on the Atoh1 function in the intestines has not been described, based on the results of studies dealing with Atoh1 and Hic1 function in the intestinal tissue and tumors, it can be assumed that Hic1 loss would probably lead to increase of *Atoh1* expression and promote tumorigenesis. Hic1 also downregulates expression of the gene encoding the transcription factor Sox9 that is required for Paneth cells maturation[89]. Sox9 depletion from the intestinal epithelium leads to loss of reserve population of intestinal stem cells (the so called +4 population) with the consequence of impaired regeneration after irradiation-induced injury[291]. In contrast, increased *Sox9* expression was observed in tumors developed in $Hic1^{+/-}$ $Apc^{+/-}$ mice, which were negative for Hic1 and exhibited accelerated growth[223]. Importantly, Sox9 has been described as a negative regulator of protein kinase C alpha (*PKCα*) in intestinal epithelial cells[59], a gene that suppresses growth of gastrointestinal tumors (reviewed in[183]), including those developed in $Apc^{+/Min}$ mice[250]. To summarize, Hic1 depletion increases *Sox9* expression, which subsequently downregulates *PKCα* expression and thus mediates tumor development. Therefore, these (and other) studies strongly suggest Sox9 function in ISCs protection from tumorigenesis, which is impaired when *Sox9* expression is imbalanced, e.g. by loss of its repressor Hic1. Hic1 has been also described as an important component of the intestinal immune system. *Hic1* is expressed in immune cells in the lamina propria and T-cell specific deletion of *Hic1* in mice results in reduced numbers of lamina propria-

resident T cells. Moreover, Hic1-deficient T-cells are not able to induce intestinal inflammation, which confirms the important role of Hic1 in intestinal tissue[34].

2 Aims of the thesis

The canonical Wnt signaling pathway is one of the most prominent signaling pathway that governs many aspects of embryonic development, participates in the maintenance of adult stem cells, and have important functions in the initiation and progression of many types of cancer. A large number of genes whose expression is dependent on the level of Wnt signaling has been discovered, however, the list is undoubtedly incomplete. Similarly, numerous genes involved in the Wnt signaling pathway regulation have been described, but the consequences of their inactivation (in the intestines) are often far from being clearly defined. *Msx1* was previously described as the Wnt target gene and its function in craniofacial development and teeth morphogenesis has been studied thoroughly. Nevertheless, the Msx1 role in cell transformation and tumor biology is still not well understood and its function in the intestines has not yet been studied at all. *Hic1* represents a well described tumor suppressor gene, however, its role in the (healthy) intestinal epithelium has not yet been fully clarified.

The aims of this thesis were to characterize the role of MSX1 transcription factor in the mammalian intestinal epithelium, especially in the context of colorectal cancer, and to study the Hic1 function in mouse intestines.

The specific aims of the theses are:

1. To characterize the Msx1 function in mouse intestinal tissue and tumors.
2. To describe the MSX1 function in human colorectal cancer.
3. To identify genes regulated by MSX1 in mouse and human intestinal cancer cells.
4. To investigate the effect of *Hic1* inactivation in the mouse intestines.

3 Materials and Methods

Experimental mice

Housing of mice and *in vivo* experiments were performed in compliance with the European Communities Council Directive of 24 November 1986 (86/609/EEC) and national and institutional guidelines. Animal care and experimental procedures were approved by the Animal Care Committee of the Institute of Molecular Genetics (Ref. 71/2014). *Apc*$^{cKO/cKO}$ mice were obtained from the Mouse Repository (National Cancer Institute, Frederick, MD); *Villin-CreERT2* and *Villin-Cre* mice[62] were kindly provided by S. Robine (Institut Curie, Centre de Recherche, Paris, France); *Msx1-LacZ* mice were obtained from the Knockout Mouse Project (KOMP) Repository; *Apc*$^{+/Min}$, *ROSA-CreERT2*, *Lgr5-EGFP-IRES-CreERT2*, *Msx1*$^{cKO/cKO}$, and immonodeficient NSG™ mice were purchased from the Jackson Laboratory (Bar Harbor, ME, US). *Hic1*$^{cKO/cKO}$ mice[271] were generated previously in our lab. Animals were housed in specific pathogen-free conditions.

Cre-mediated gene recombination

For expression profiling, adult mice were gavaged with 5 mg of tamoxifen (Sigma-Aldrich); 0.3 mg of tamoxifen was used in the survival experiment; 1 mg of tamoxifen was used in all other experiments. Tamoxifen was dissolved in ethanol (100 mg/ml) and prior to gavage combined with mineral oil (1:1 ratio). Mice of 6 weeks or older were used in all experiments. Mice were sacrificed by cervical dislocation at various timepoints after a single dose (100 µl) of tamoxifen/oil mixture. Intestines were dissected, washed in phosphate-buffered saline (PBS), fixed in 4% (v/v) formaldehyde (FA, Sigma-Aldrich) in PBS overnight, embedded in paraffin, sectioned, and stained.

DSS treatment

Msx1$^{cKO/cKO}$ *Villin-Cre* mice were administered with 2% (w/v) DSS (MP Biomedicals; MW36–50 kDa) in drinking water for 5 days to induce damage in the colon. Colons were collected upon DSS withdrawal at day 2 (acute colitis), day 5 (beginning of the regenerative phase), and day 9 (late regenerative phase). Two Cre^{+} and two Cre^{-} littermates were used; untreated mice of the same genotype were used as a control.

Total body irradiation

Msx1$^{cKO/cKO}$ *Villin-Cre* mice were exposed to single total body irradiation of 5 Gy during 5 minutes. The X-RAD 225 XL instrument (PRECISION X-Ray Inc) with 0.5 mm Cu (F5) filter was used; voltage and current of the X-ray tube was set to 225 kV/13.3 mA. Mice were sacrificed 2, 4, and 6 days after irradiation and the small intestines were collected for each timepoint. Two Cre^{+} and two Cre^{-} littermates were used; untreated mice of the same genotype were used as a control.

Xenotransplantation

SW620 single cell clones with the *MSX1* gene disrupted (n = 3) and control cells with intact *MSX1* (n = 3) were cultured to 90-100% confluency, harvested, and resuspended in PBS. NSG^{TM} mice were injected with 1×10^{7} cells (in 100 µl PBS) into the lumbar back area. Mice were sacrificed 28 days after injection, tumors were resected and weighed.

Crypt isolation, organoid culture, and 4-hydroxytamoxifen treatment

Mouse intestinal crypts were isolated according to previously published protocols[309, 310]. Briefly, approximately 5 cm of the proximal jejunum was dissected and washed from inside with cold PBS, cut opened, and the villi were scraped off using a coverslip. The tissue was transferred into 50ml Falcon tube with 15ml of ice-cold PBS and vigorously shaked. The PBS with debris was removed, fresh PBS was added and the process of shaking was repeated for 5-10 times, until the PBS after shaking was clear. The intestines were incubated in 5mM ethylene-diamine-tetra-acetic acid (EDTA, Merck) in PBS at 4°C for 30 min on a rocking platform. The PBS with EDTA was removed, 7-8 ml of fresh PBS were added, and crypts were released by gentle shaking. The suspension was passed through 70µm strainer (Corning) and centrifuged at 300g at 4°C for 5min. The pellet was washed with 5ml ice-cold Advanced Dulbecco's Modified Eagle's Medium (DMEM)/F-12 medium (Gibco), resuspended in Matrigel (BD Biosciences), and small drops of the mixture were placed to pre-warmed 24-well plate. The plate was incubated for 10 min at 37°C and a complete organoid growth medium [Advanced DMEM/F-12 supplemented with 10mM HEPES, 1× Glutamax, 1mM N-acetyl-cysteine, 1x Penicilin/Streptomycin, 1× N-2 supplement, 1× B-27 supplement (all from Thermo Fisher Scientific), 10% Rspo1-conditioned media, 10% Noggin-conditioned media, mEGF (50 ng/ml; Peprotech), and Primocin (100 µg/ml, Invivogen)] was poured to the embedded crypts. A fresh complete organoid growth medium was given to the organoid culture every 2-3 days. In growth

media for colon organoids, Wnt3a-conditioned medium was added to final 50% concentration. Organoids were passaged every 5-7 days by mechanical disruption using 200-µl plastic pipette tip. For Cre-mediated recombination, organoids were treated with 1µM 4-OHT (Sigma-Aldrich); control organoids were treated with same volume of ethanol. Alternatively, single-cell suspension of crypt cells was obtained by incubation of freshly isolated crypts with Dispase (100 mg/ml, dilution 1:200 in serum free DMEM; Gibco) in a mixing block for 5 min at 37°C and 800 rpm for three times. The suspension was diluted in DMEM supplemented with 10% fetal bovine serum (FBS; Gibco) and passed through a 40-µm strainer (352340, Corning) to obtain single cells.

Fluorescence-activated cell sorting (FACS)

Single cell suspension obtained from the *Lgr5-EGFP-IRES-CreERT2* small intestinal crypts was stained with following antibodies: Fluorescein isothiocyanate (FITC)-conjugated anti-CD45 (ExBio), Allophycocyanin (APC)-conjugated anti-Epithelial cell adhesion molecule (EpCam, eBioscience) and Phycoerythrin (PE)-conjugated anti-CD24 (eBioscience) for 30 min at 4°C. Single cells were gated by forward scatter (FSC), side scatter (SSC), and negative staining for Hoechst 33258 (Sigma-Aldrich). $CD45^-EpCAM^+$ population of epithelial cells was further sorted by GFP expression to $CD24^+/GFP^-$ (Paneth cells) and $CD24^+/GFP^+$. According to the SSC pattern, $CD24^+/GFP^+$ population was further divided into $CD24^+/GFP^+$ large cells ("intermediate" cells) and $CD24^+/GFP^+$ small cells (stem cells). Alternatively, epithelial cells obtained from $Apc^{cKO/cKO}$ *Villin-CreERT2* small intestine 7 days after tamoxifen administration were processed as described previously[368]. Briefly, cells were washed with RNase-free staining buffer [SB; PBS supplemented with 1% RNase-free bovine serum albumin (BSA, Gemini Bioproducts) and 0.0025% RNasin Plus (Promega)] and incubated with Fixable Viability Dye eFluor 780 (Thermo Fisher Scientific) and APC-EpCam antibody. Cells were fixed in 4% paraformaldehyde (PFA; Electron Microscopy Sciences) diluted in PBS for 15 min at 4°C and permeabilized in SB buffer with 0.1% Triton X-100 (Sigma-Aldrich) for 10 min at 4°C on a rocking platform. After fixation, the anti-Msx1 antibody (R&D Systems) was used followed by Alexa Fluor 488-conjugated rabbit anti-goat IgG (Thermo Fisher Scientific). $eFluor\ 780^+/EpCAM^+$ single cells were sorted as Msx1 (highly) positive (+) or negative (-). Sorting was performed on Influx cell sorter (BD Biosciences).

Cell cycle analysis and proliferation assay

Cell number and viability were determined using the automated cell counter Cellometer Auto T4 (Nexcelom Bioscience) based on the trypan blue exclusion method or by CellTiter-Blue® reagent (Promega) and EnVision® Multilabel Reader (PerkinElmer). Cells were harvested and washed with ice-cold PBS, fixed in 70% ethanol, dyed with propidium iodide (PI; Sigma-Aldrich) and the cell cycle was analyzed using a BD FACSCanto™ II flow cytometer (BD Biosciences) and FlowJo™ software.

Microarray analysis

Total RNA was isolated from *Apc*$^{cKO/cKO}$ *Villin-CreERT2* intestinal epithelium 2 and 4 days after administration of 5 mg of tamoxifen by gavage; mice administered with the solvent (ethanol and mineral oil mixture) were used as controls. Four biological replicates were used for each timepoint. The RNA samples were analyzed using MouseRef-8 v2.0 Expression BeadChip (Illumina). Raw data were processed using the beadarray package of Bioconductor and analyzed as described previously[215]. Gene set enrichment analysis (GSEA) was performed using the Enrichr gene analysis tool[123, 181]. Alternatively, total RNA was isolated from the *Apc*$^{cKO/cKO}$ and *Apc*$^{cKO/cKO}$*Msx1*$^{cKO/cKO}$ mouse small intestinal or colonic epithelium 7 days after administration of 1 mg of tamoxifen by gavage; mice administered with the solvent were used as controls. Four biological replicates were used for both mouse strains. RNA samples obtained from the small intestine were processed and analyzed as described above. RNA samples obtained from the colon were amplified and labeled using GeneChip WT PLUS Reagent Kit (Applied Biosystems) following the supplier's protocol and starting with 250 ng of total RNA. Labeled single-stranded DNA was hybridized onto GeneChip Mouse Gene 2.0 ST arrays (Affymetrix) using GeneChip Hybridization, Wash, and Stain Kit (Applied Biosystems) following the supplier's protocol. Arrays were scanned using GeneChip 3000 7G Scanner (Affymetrix). Total RNA isolated from SW620 cell clones with the *MSX1* gene disrupted (n = 8) or intact (n = 4) was used. The RNA samples were analyzed using Human HT expression BeadChip V4 (Illumina). Raw data were processed and analyzed as described above. The quality of isolated mRNA was checked using Agilent Bioanalyzer 2100; RNAs with RNA integrity number (RIN) above 8 were further processed. The microarray analysis was performed in the Genomics and Bioinformatics facility, IMG, Prague, Czech Republic.

Human samples

All methods used to collect the human specimens were performed in accordance with the relevant national and EU guidelines and regulations. The study was approved by the Ethics Committee of the Third Faculty of Medicine, Charles University in Prague. Informed consents have were from all patients participating in the study. Paired samples of normal and neoplastic colonic tissue were obtained from patients undergoing either polypectomy of colonic adenomas or surgical resection of sporadic CRC (patient data are summarized in Table 1). The tumor and corresponding normal colonic mucosa samples were immediately frozen and stored in liquid nitrogen. None of the patients underwent radiotherapy or chemotherapy before operation. Samples were processed as described previously[147]. Briefly, human samples were disrupted in 600 µl of lysis buffer by green ceramics beads and MagNA Lyser Instrument (Roche Life Science) and total RNA was extracted using RNeasy Mini kit (Qiagen) according to manufacturer's instructions. Complementary DNA (cDNA) synthesis was performed in 20-µl reaction using 1 µg of total RNA, random hexamers and RevertAid reverse transcriptase (Thermo Fisher Scientific) according to manufacturer's protocol. PCR reactions were run in triplicates using LightCycler 480 Probes Master and Universal Probe Library (UPL) hydrolysis probes and LightCycler 480 Instrument (all from Roche Life Sciences). The primer pairs and corresponding UPL probes are listed in Table 2. Threshold cycle (Ct) values for each triplicate were normalized by geometric average of housekeeping genes ubiquitin B (*UBB*) and TATA-box binding protein (*TBP*). Resulting values were averaged to obtain ΔCt values for biological replicates. Relative mRNA abundance (ΔCt in healthy tissue − ΔCt in neoplastic tissue) was correlated with the histological grade of tumor samples using the rank-order Spearman's (ρ) and Kendall's (τ) coefficient.

Table 1 | Patient summary and histopathological grade of colorectal neoplasia specimens.
The *MSX1* and *SOX17* gene expression levels were analyzed in the total amount of 72 neoplasia specimens and compared to the gene expression level in the matching healthy mucosa.

	Age (median, min.-max.)	Gender (M/F)
HYP	69 (59-77)	6/3
LGD	67 (53-89)	12/15
HGD	64 (36-85)	15/9
CRC	82 (63-90)	6/6

Cell culture

HEK293, SW480, and SW620 cell lines were purchased from the American Type Culture Collection (Cat. No.: CRL-1573, CCL-228, and CCL-227). STF cells[411] were kindly

provided by Q. Xu and J. Nathans (Johns Hopkins University, Baltimore, MD). HEK293 and STF cells were maintained in DMEM supplemented with 10% FBS, penicillin, streptomycin, and gentamicin (all antibiotics were purchased from Invitrogen). SW480 and SW620 cells were maintained in Iscove's Modified Dulbecco's Medium (IMDM) supplemented with 10% FBS, penicillin, streptomycin, gentamicin, NEA (Gibco), and Glutamax (Gibco). Conditioned media for organoid cultures were obtained from cells producing mouse R-spondin1 or mouse Noggin (dilution 1:10; cells were kindly donated by M. Maurice, University Medical Center Utrecht, Utrecht, The Netherlands). For Wnt signaling activation, HEK293 cells were treated with GSK3β inhibitor BIO [Sigma-Aldrich; final concentration 1μM; the stock solution was prepared in dimethyl sulfoxide (DMSO); control cells were treated with solvent only] or by conditioned media (CM) obtained from cells producing the mouse Wnt3a ligand (dilution 1:1; cells were kindly donated by M. Maurice, University Medical Center Utrecht, Utrecht, The Netherlands); control cells were treated with the same dilution of CM obtained from cells non-producing the Wnt3a ligand. Both treatments were performed overnight.

Cell viability test

Cells were seeded to 96-well Assay Plates (Corning) at approximately 12.5% confluency in 100 μl of IMDM. Six hours after seeding, 10 μl of alamarBlue reagent (Thermo Fisher Scientific) was added to each well. After 60 min, the fluorescence intensity was measured in the culture medium using EnVision® Multilabel Reader (PerkinElmer). The measurement was repeated every 24 hours for 5 days.

RNA purification, cDNA synthesis, quantitative real-time polymerase chain reaction (qRT-PCR)

Total RNA from cell lines and mouse tumors was isolated using TRI Reagent (Sigma-Aldrich), total RNA from epithelial or crypt cells, sorted cells and organoids was isolated by RNeasy Micro Kit or RNeasy Mini Kit (both kits were purchased from Qiagen), total RNA from fixed sorted cells was isolated using miRNeasy FFPE Kit (Qiagen); RNA isolation was performed according to the manufacturer's protocol. Complementary DNA synthesis was performed in 20-μl reaction using random hexamers and 1 μg of total RNA (or the whole eluate when the RNeasy Micro Kit or miRNeasy FFPE Kit was used). RNA was reverse transcribed using RevertAid Reverse Transcriptase or MAXIMA Reverse Transcriptase (both were purchased from Thermo Fisher Scientific) following

manufacturer's protocol; for cDNA synthesis from fixed sorted cells, the QuantiTect kit (Qiagen) was used. Quantitative RT-PCR reactions was performed in triplicates using SYBR Green I Master Mix and measured by LightCycler 480 apparatus (both from Roche Life Science). For a list of primers, see Table 2.

Immunoblotting

The immunoblotting procedure was performed as described in a detailed protocol previously[197]. Briefly, cultured cell were incubated in lysis buffer [50mM Tris (pH 7.8), 400mM NaCl, 0.5% Triton X-100] supplemented with the protease inhibitor cocktail (Roche Life Science) for 30 min at 4°C. The lysates were centrifuged at 20000g for 15 min at 4°C. Protein concentration in the lysate was assessed using Bradford Reagent (VWR International) and EnVision® Multilabel Reader (PerkinElmer). The lysate was mixed with Laemmli Sample Buffer (Merck), incubated at 95°C for 5 min, separated by sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE), and transferred to a nitrocellulose membrane using Trans-Blot® SD semi-dry transfer cell (Bio-Rad). The membranes were blocked using 5% (w/v) non-fat milk with 0.05% Tween-20 (Merck) in PBS and incubated with primary antibodies overnight at 4°C. Membranes were washed in PBS with 0.05% Tween-20 and incubated for 1 hour with horseradish peroxidase (HRP)-conjugated secondary antibody. The HRP activity was detected by Pierce ECL Western Blotting Substrate (Thermo Fisher Scientific). Primary antibodies are listed in Table 3, peroxidase-conjugated anti-goat, anti-mouse, and anti-rabbit secondary antibodies were purchased from Sigma-Aldrich.

Immunocytochemistry

SW620 EGFP-MSX1 cells were stained as described previously[178]. Briefly, cells were grown on round coverslips in wells of 24-well plate, fixed in 4% PFA diluted in PBS for 20 min at room temperature (RT), and treated with 0.25% Triton X-100 in PBS for 10 min. Cells were incubated with primary antibody diluted in DMEM supplemented with 10% FBS at 4°C overnight. Subsequently, secondary antibodies diluted in DMEM/FBS were added to cells for 1 hour incubation at RT. Cells were counterstained with hematoxylin (Vector Laboratories) or DAPI nuclear stain (Sigma-Aldrich). Primary antibodies are listed in Table 3; goat-anti-rabbit and rabbit-anti-goat secondary antibodies conjugated with biotin or Alexa Fluor 488 or 594 were purchased from Thermo Fisher Scientific.

Immunohistochemistry, β-galactosidase and periodic acid-Schiff staining

The tissue was fixed in 4% (v/v) formaldehyde (FA; Sigma-Aldrich) in PBS overnight and embedded in paraffin using an automatic tissue processor (Leica). Paraffin embedded tissue was cut to 6-µm sections that were incubated at 42°C overnight and dewaxed in xylene. Antigen retrieval was performed in 10mM citrate buffer (pH 6.0) in a steam bath for 20 min. Endogenous peroxidase was blocked by incubation in 0.2% H_2O_2 (Sigma-Aldrich) diluted in methanol (Merck, Kenilworth, NJ, USA) for 25 min. Slides were incubated with primary antibody diluted in Tris-buffered saline (TBS) buffer supplemented with 5% serum (Vector Laboratories) and 1% BSA (Sigma-Aldrich) at 4°C overnight. Subsequently, biotin-conjugated secondary antibodies diluted in TBS supplemented with 5% serum were added to slides for 1 hour incubation at RT. The signal was enhanced by 30 min incubation with Vectastain ABC kit (Vector Laboratories) and developed in 3,30-diaminobenzidine (DAB) solution (Vector Laboratories; 30 mg dissolved in 90 ml of 50mM Tris, pH 7.5). Sections were counterstained with hematoxylin (PENTA), dehydrated and assembled in Solakryl BMX (Draslovka Holding B.V.). Primary antibodies are listed in Table 3. Biotin-conjugated anti-goat, anti-mouse, and anti-rabbit secondary antibodies and Alexa Fluor 488-conjugated anti-mouse and anti-rabbit secondary antibodies were purchased from Thermo Fisher Scientific. *Apc*$^{cKO/cKO}$ *Msx1*$^{wt/LacZ}$ *Villin-CreERT2* mice were sacrificed 7 days after administration of a single dose of tamoxifen; mice administered with the solvent were used as controls. *LacZ*-expressing tissue was pre-fixed in ice-cold fixative [1% PFA, 0.2% glutaraldehyde (Merck), 0.02% NP40 in PBS] at 4°C for 1 hour on a rocking platform. The fixative was removed and the tissue was washed three times in PBS for 20 min at RT on a rocking platform. The β-galactosidase substrate solution [5mM $K_3Fe(CN)_6$, 5mM $K_4Fe(CN)_6$, 2mM $MgCl_2$, 0.02% NP40, 0.01% sodium deoxycholate, and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal, 1 mg/ml; Sigma-Aldrich) in PBS] was added and the tissue was incubated in the dark overnight. The substrate was removed and the tissue was washed twice in PBS for 20 min at RT on a rocking platform and fixed in 4% FA at 4°C overnight. After fixation, the tissue was embedded in paraffin and processed as described above. Sections were counterstained with the nuclear fast red stain (Sigma-Aldrich). The periodic acid-Schiff (PAS) staining was performed using the P.A.S. acc. Hotchkiss-McManus kit (DiaPath) according to the manufacturer's instructions.

In situ hybridization (ISH)

Paraffin embedded tissue was cut to 8-µm sections that were incubated at 42°C for 1 hour and dewaxed as described above. Sections were treated with 0.2M HCl for 15 min at RT and incubated with Proteinase K (EO0491, Thermo Fisher Scientific) diluted in PBS (concentration 30 µg/1 ml) for 15 min at RT. Sections were rinsed with 0.2% Glycine (PENTA) dissolved in PBS, fixed with 4% (v/v) PFA in PBS for 10 min at RT, treated with 0.25% acetic anhydride (Sigma-Aldrich) diluted in 0.01M triethanolamine hydrochloride solution (Sigma-Aldrich) with pH 8.0 for 10 min. Digoxigenin-labeled probes were diluted in hybridization solution consisting of 50% formamide (Amresco), 5× Saline-Sodium Citrate buffer (SSC; 20× concentrated SSC contains 3M NaCl and 0.3M sodium citrate, pH 4.5), 2% Blocking Reagent (11096176001; Roche Life Science), 5mM EDTA, 0.05% CHAPS (Sigma-Aldrich), 50 µg/ml heparin (Sigma-Aldrich), and 1 µg/ml yeast total RNA (Sigma-Aldrich) and incubated on slides for 48-72 hours at 65°C. Sections were washed in 50% formamide diluted in SSC (pH 4.5) with 0.1% Tween-20 at 62°C. Hybridized slides were developed using alkaline phosphatase-conjugated sheep anti-digoxigenin Fab antibody (Roche Life Science) diluted in MATB buffer [100mM maleic acid (pH 7.5), 150mM NaCl, 0.1% Tween-20] with 0.5% Blocking Reagent at 4°C overnight. The signal was detected using a mixture of nitro-blue tetrazolium chloride [NBT; 100 mg/ml in 70% dimethylformamide (DMF)] and 5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt (BCIP; 50 mg/ml in 100% DMF) substrates [17µl each in 1 ml of buffer consisting of 0.1M NaCl, 0.1M Tris (pH 9.5), 0.05M $MgCl_2$] for 1-6 hours at RT. Sections were mounted in Mowiol (Sigma-Aldrich).

Synthesis of ISH probes

Mouse *Msx1* and *Sox17* cDNAs were purchased from Addgene (#21024 and #50781). *Msx1* cDNA was directly cut out by restriction enzymes from the donor vector and ligated into pBluescript KS II (Stratagene) vector; *Sox17* cDNA was PCR amplified from the donor vector and ligated into pBluescript KS II. Mouse *Ascl2* and *Olfm4* cDNAs were PCR amplified from cDNA reverse transcribed from total RNA isolated from wild-type mouse small intestinal epithelial cells and cloned into pBluescript KS II (primers used for PCR amplification are listed in Table 2). Plasmids were linearized and digoxigenin-labeled RNA probes were synthesized using DIG RNA Labeling Mix (Roche Life Science) and T7 or T3 RNA polymerase (Thermo Fisher Scientific) for sense and antisense probes, respectively. Probes were purified using mini Quick Spin RNA Columns (Roche Life Science).

RNA interference (RNAi)

Cells were transfected with 10nM small interfering RNAs (siRNAs) targeting the *CTNNB1* (β-catenin) gene (s437; Ambion) or control siRNAs (D001206-13-20; Dharmacon) using Lipofectamine RNAiMax (Invitrogen) according to manufacturer's protocol. Cells were re-transfected 2 days after the first transfection to increase the effect of RNA interference and harvested 2 days after the second transfection.

Transfection, Lentivirus production and purification

To produce lentiviral particles, one 10 cm Petri dish with HEK293FT cells (Invitrogen) was seeded at ~30% confluency one day before transfection in IMDM media. The transfection was performed using Lipofectamine® 2000 (Thermo Fisher Scientific) in serum-free OptiMEM medium (Thermo Fisher Scientific) according to manufacturer's protocol. For one 10 cm Petri dish, following amounts of plasmids were used: 4 µg psPAX2 (Addgene #12260), 2.7 µg pCMV-VSV-G (#8454), and 5.3 µg lentiCas9-Blast (#52962). After 6 hours of incubation with plasmid DNA at 37°C, media was changed to fresh IMDM. After 48 hours, media was removed and centrifuged at 3000g at 4°C for 15 min to spin down cell debris. The supernatant (~ 12 ml) was transferred to a new plastic tube, 25 % of the volume (~ 3 ml) of PEGit Virus Precipitation Solution (System Biosciences) was added and the mixture was manually rotated 10-20 times. The mixture was incubated at 4°C overnight and then centrifuged at 1500g at 4°C for 30 min to spin down lentiviral particles. The supernatant was discarded and the pellet was gently resuspended in 150 µl of PBS. Generation of lentiCRISPR viruses was analogous, lentiGuide-Puro plasmids (#52963) encoding corresponding guide RNAs (gRNAs) were used.

Disruption of the human *APC* gene

Exon 15 of the *APC* gene was targeted in STF cells[411] using the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) system. Three different gRNAs were cloned into the lentiCRISPRv2 vector (Addgene, #52961) as described previously[304, 319]. Guide RNAs were designed using CRISPR Design Tool available at crispr.mit.edu; the list of gRNA sequences is provided in Table 2. Cells were co-transfected with lentiCRISPRv2 plasmid and pARv-RFP reporter[158] (Addgene #60021) containing the appropriate gDNA sequence recognized by gRNA. RFP^+ cells were sorted into 96-well plates and expanded as single cell clones. Control cells were transduced with the empty (BsmBI digested and self-ligated) lentiCRISPRv2 vector and processed in

an analogous way. Generation of STF cells harboring truncation in exon 10 of the *APC* gene was described previously[372]. The scheme of the targeted locus can be found in Figure 17.

Tagging and disruption of the *MSX1* gene in human cells

To generate cells producing MSX1 protein fused at its N-terminus with EGFP, SW620 cells were co-transfected with the lentiCRISPRv2 vector containing gRNA targeting exon 1 of the *MSX1* gene together with a synthetic template (purchased from GenScript) for homology-directed repair of the *MSX1* locus and corresponding pARv-RFP reporter vector (Addgene #60021). RFP^+ cells were sorted into 96-well plates and expanded as single cell clones. Individual clones were PCR screened for the presence of correct insertion of the exogenous template, genotyping primers are provided in Table 2. The scheme of the targeted locus can be found in Figure 45. To disrupt the *MSX1* gene, SW620 cells stably producing the Cas9 enzyme were generated by retroviral transduction using the lentiCas9-Blast vector (#52962, Addgene) and batch-selected using 10 μg/ml blasticidin (Gibco). To introduce truncations in the *MSX1* gene, SW620/Cas9-Blast cells were transduced with lentiGuide-Puro (#52963, Addgene) lentiviral particles targeting the first exon of the *MSX1* gene and selected with 6 μg/ml puromycin (Thermo Fisher Scientific). The transduction was performed in four technical replicates, two lentiGuide-Puro plasmids targeting different sites in the *MSX1* first exon were used; cells transduced with empty (BsmBI digested and selfligated) lentiGuide-Puro plasmid were used as a control. Correct targeting was verified by PCR amplification and sequencing. Selected polyclonal cultures were analyzed by Western blotting. The scheme of the targeted locus can be found in Figure 41.

Chromatin immunoprecipitation (ChIP)

SW620 cells (10 million cells per ChIP) producing EGFP-MSX1 fusion protein were crosslinked by 1% FA, lysed using HighCell ChIP kit (C01010062, Diagenode), and the isolated chromatin was sheared by sonication with three runs of 10 cycles (30 s ON/30 s OFF) at high power setting using Bioruptor sonicator (Diagenode). The sheared chromatin from SW620 GFP-MSX1 or parental cells was incubated with GFP-Trap®_MA (Chromotek) magnetic beads and further processed as described in the protocol provided by the manufacturer. Antibodies against H3K4me3 (C15410003) and negative control rabbit IgG were purchased from Diagenode. After decrosslinking, the DNA was isolated using QIAquick PCR purification kit (Qiagen) and quantitative PCR reactions were

performed using a LightCycler 480 Instrument. After decrosslinking, DNA was isolated using QIAquick PCR purification kit (Qiagen). Occupancy of the *ASCL2* and *SP5* gene regulatory regions by MSX1 was assayed by PCR; primers are listed in Table 2. The recovery of the particular promoter locus was calculated as the relative amount of immunoprecipitated DNA compared to input DNA (percentage of input).

Luciferase reporter assay

The luciferase reporter assay was performed using Dual-Glo Luciferase Assay System and GloMax® 20/20 Luminometer (all from Promega), according to the manufacturer's protocol. To test the human *SP5* promoter activity, a region of the promoter containing TCF/β-catenin and MSX1 binding sites was PCR-amplified from human genomic DNA and cloned into the pTA-Luc vector (Invitrogen). To test the mouse *Sox17* promoter activity, a region of the promoter containing two Msx1 binding sites was PCR-amplified from mouse genomic DNA and cloned into pGL4.26 vector (Promega, # 9PIE844). The scheme of the PCR-amplified *SP5* and *Sox17* locus can be found in Figure 46 and 47, respectively. Primers are listed in supplementary Table 2. All luciferase assays were performed in triplicates; three MSX1-deficient and three control clones of SW620 cells were used. Luciferase activity was normalized to Renilla.

Wound healing assay

SW620 cells with *MSX1* gene disrupted (n = 3) and intact (n = 2) were cultured to 100% confluency and treated with mitomycin C (Sigma-Aldrich) diluted to final concentration 10 µg/ml in IMDM. After 2 hours, several scratches ("wounds") were made in the cell layer using a 100µl plastic pipette tip and the medium was changed for fresh IMDM. Microscopic pictures of six sites from the scratched area of each cell clone were taken at 0 hours and 22 hours. The scraped area was marked manually using the FiJi software. The percentage of "healed" area was calculated from the difference of the scraped area at 22 hours and at 0 hours.

Statistical analysis of data

The results of the qRT-PCR analysis were evaluated by the Student's t-test. The relative mRNA abundance (ΔCt in healthy tissue – ΔCt in neoplastic tissue) was correlated with the histological grade of tumor samples using the rank-order Spearman's (ρ) and Kendall's (τ) coefficient. Datasets obtained using RNA microarrays were analyzed within the oligo

and limma packages of Bioconductor[40, 88, 338]. Moderated t-test was used to detect differentially expressed genes (DEGs) between experimental groups: at least two-fold change difference in gene expression and Storey's q-value[348] less than 0.05 were considered significant. The statistic of the survival (Kaplan-Meier) curves was calculated using the log rank test[23] and publically available online calculator (http://astatsa.com/LogRankTest/). The statistical analysis of data obtained from microarray analysis was performed by H. Strnad and M. Kolar from the Genomics and Bioinformatics facility, IMG, Prague, Czech Republic.

Raw expression data repository

Minimum Information About a Microarray Experiment (MIAME) compliant data were deposited to the ArrayExpress database (E-MTAB-6915, E-MTAB-6930, E-MTAB-6928, and E-MTAB-6909).

Table 2 | List of DNA oligonucleotides used in the study.

List of primers used for qRT-PCR

Gene	Organism	Primer sequence (5′ to 3′)	
β-actin	Mouse	Forward:	GATCTGGCACCACACCTTCT
		Reverse:	GGGGTGTTGAAGGTCTCAAA
β-actin	Human	Forward:	GGCATCCTCACCCTGAAGTA
		Reverse:	AGGTGTGGTGCCAGATTTTC
ABHD12B	Human	Forward:	CGGAAGAAAAATTGCTGCTC
		Reverse:	TCACCCCAGGTTCAACTCTC
Ascl2	Mouse	Forward:	AAGCACACCTTGACTGGTACG
		Reverse:	AAGTGGACGTTTGCACCTTCA
ASCL2	Human	Forward:	GCGAGCTACTCGACTTCTCC
		Reverse:	CTCGGCTTCCGGGGCTGAGG
Axin2	Mouse	Forward:	TAGGCGGAATGAAGATGGAC
		Reverse:	CTGGTCACCCAACAAGGAGT
AXIN2	Human	Forward:	TGAGGTCCACGGAAACTGTTGACAGT
		Reverse:	CCCTCCCGCGAATTGAGTGTGA
Bves	Mouse	Forward:	GAACTGGCGAGAGATTCACC
		Reverse:	ATCATCACATCCAAGGCACA
CDX2	Human	Forward:	TGCTGCAAACGCTCAACCCCGG
		Reverse:	CGGCTTTCCTCCGGATGGTGATG
CHGA	Mouse	Forward:	GCGCCGGGCAAGTTTTTGCC
		Reverse:	GGGCTGGGTTTGGACAGCGAG
Cryptdins	Mouse	Forward:	AGGAGCAGCCAGGAGAAG
		Reverse:	ATGTTCAGCGACAGCAGAG
CTSZ	Human	Forward:	GCTTCTGCTGCTCGTGCT
		Reverse:	GTTGACACCATCCACATTGC
CTNNB1 (β-catenin)	Human	Forward:	TTCCAGACACGCTATCATGC
		Reverse:	AATCCACTGGTGAACCAAGC
DEPDC7	Human	Forward:	ACCTAAGAGGCAGTCCACCA
		Reverse:	GTCTGGTTGCTCAGGAAAGC
EGFP	Aequorea victoria	Forward:	GACGTAAACGGCCACAAGTT
		Reverse:	GAACTTCAGGGTCAGCTTGC
ENTPD8	Human	Forward:	AGCGTCTAAGCACAGCTTCC
		Reverse:	TCCACCAGGAGGAGAATGAG
GAPDH	Mouse	Forward:	AACTTTGGCATTGTGGAAGG
		Reverse:	ATCCACAGTCTTCTGGGTGG
KRT23	Human	Forward:	GCCTCCGAAGGACCTTAGAC
		Reverse:	AGATCTTCCCTGGGACCTGT
Lgr5	Mouse	Forward:	CCTGTCCAGGCTTTCAGAAG
		Reverse:	CTGTGGAGTCCATCAAAGCA
Lysozyme	Mouse	Forward:	CCTGACTCTGGGACTCCTCCTGCT
		Reverse:	CTAAACACACCCAGTCGGCCAGGC
Mdga1	Mouse	Forward:	CCTCACACCCTACACCACCT
		Reverse:	GGGCCCAGTATTAGGAGAGC
Msx1	Mouse	Forward:	CTCTCGGCCATTTCTCAGTC
		Reverse:	TTGGTCTTGTGCTTGCGTAG
MSX1	Human	Forward:	AGAAGATGCGCTCGTCAAAG
		Reverse:	GGCTTACGGTTCGTCTTGTG

Msx2	Mouse	Forward:	AATTCCGAAGACGGAGCAC
		Reverse:	CGGTTGGTCTTGTGTTTCCT
MSX2	Human	Forward:	CGGTCAAGTCGGAAAATTCA
		Reverse:	GAGGAGCTGGGATGTGGTAA
Mtus2	Mouse	Forward:	TCGTCCTCCTGGCTATTCAC
		Reverse:	CCCTTGGGTGTGTCCTTAGA
Mylk3	Mouse	Forward:	CCCAGGAAGAACTGAAGCTG
		Reverse:	CGACCCCCTCCTAAGACTTC
Nkd1	Mouse	Forward:	AGGACGACTTCCCCCTAGAA
		Reverse:	TGCAGCAAGCTGGTAATGTC
NKD1	Human	Forward:	GCCTCCTGAGAAGACTGACG
		Reverse:	TTGCCGTTGTTGTCAAAGTC
Olfm4	Mouse	Forward:	TGGGCAGAAGGTGGGACTGTGT
		Reverse:	TGTCAGCGGGAAAGGCGGTA
PCNA	Mouse	Forward:	CCACATTGGAGATGCTGTTG
		Reverse:	CAGTGGAGTGGCTTTTGTGA
RASL10B	Human	Forward:	GGGGAGCCCCTACTTCTCTC
		Reverse:	ACCGTCAGGACCAACCATTG
RASL11B	Human	Forward:	CCTGGCTCTTCAGGTTCAAG
		Reverse:	GTGGAGCTGGCTGATGAGTT
Slc5a6	Mouse	Forward:	GCCCTAGGAATTGTCTGCAA
		Reverse:	GGCAAGGGAACACTGCATAG
SORBS2	Human	Forward:	AATTCACATGGTGCACAAGC
		Reverse:	AGACCGATCTCTTGGTCGAA
Sox17	Mouse	Forward:	AAGAAACCCTAAACACAAACAGCG
		Reverse:	TTTGTGGGAAGTGGGATCAAGAC
SP5	Mouse	Forward:	GGACAGGAAACTGGGTCGTA
		Reverse:	AATCGGGCCTAGCAAAAACT
SP5	Human	Forward:	ACTTTGCGCAGTACCAGAGC
		Reverse:	ACGTCTTCCCGTACACCTTG
Stk32b	Mouse	Forward:	GTGCAGAAGCGAGACACAAA
		Reverse:	CTGTAGGTGGTAGCGCAGGT
Sucrose Isomaltase	Mouse	Forward:	TTCAAGAAATCACAACATTCAATTTACCTAG
		Reverse:	CTAAAACTTTCTTTGACATTTGAGCAA
TMEM47	Human	Forward:	TGCCATCATTCTCATTGCAT
		Reverse:	AACCCCAGTTGAACTCATGG
TNFRSF19 (Troy)	Mouse	Forward:	GCTCAGGATGCTCAAAGGAC
		Reverse:	CCAGACACCAAGACTGCTCA
TNFRSF19 (TROY)	Human	Forward:	CTATGGGGAGGATGCACAGT
		Reverse:	TCTCCACAAGGCACACACTC
Trpm	Mouse	Forward:	GGCACACAGAGTGGACTTGA
		Reverse:	AAGCCACGAAAATCTGATCG
Ttn	Mouse	Forward:	CCTGCCTCAGTGGAAGAGAC
		Reverse:	TTCTGGCTCTGGTTCCAGTT
Ubb	Mouse	Forward:	ATGTGAAGGCCAAGATCCAG
		Reverse:	TAATAGCCACCCCTCAGACG
UBB	Human	Forward:	GCTTTGTTGGGTGAGCTTGT
		Reverse:	TCACGAAGATCTGCATTTTGA

List of primers and UPL probes used for qRT-PCR

Gene	Organism	Primer sequence (5´ to 3´)		UPL probe number
MSX1	Human	Forward:	CTCGTCAAAGCCGAGAGC	70
		Reverse:	CGGTTCGTCTTGTGTTTGC	
MSX2	Human	Forward:	TCGGAAAATTCAGAAGATGGA	70
		Reverse:	CAGGTGGTAGGGCTCATATGTC	
SOX17	Human	Forward:	ACGCCGAGTTGAGCAAGA	61
		Reverse:	TCTGCCTCCTCCACGAAG	
TBP	Human	Forward:	GAACATCATGGATCAGAACAACA	87
		Reverse:	ATAGGGATTCCGGGAGTCAT	
UBB	Human	Forward:	TCACATTTTCGATGGTGTCACT	39
		Reverse:	TCACATTTTCGATGGTGTCACT	

List of primers used for cDNA amplification for subsequent cloning into pBluescript KS II vector (designed for synthesis of probes for *in situ* hybridization)

Gene	Organism	Primer sequence (5´ to 3´)	
Ascl2	Mouse	Forward:	AGTGGATCCATGGAAGCACACCTTGACTG
		Reverse:	GAGGTCGACTCAGTAGCCCCCTAACCAAC
Olfm4	Mouse	Forward:	GCTATGGCCAAGGAGGTGGT
		Reverse:	TGCTCTGAATTCTTTCCTGCATC
Sox17	Mouse	Forward:	CGCTCTAGAATGAGCAGCCCGGATGCGGGA
		Reverse:	GAGGTCGACCTAGCATCTTGCTTAGCTCTG

List of oligos used for cloning into lentiCRISPR and pARv-RFP vectors

Gene	Organism	Primer sequence (5´ to 3´)		Recipient plasmid
APC	Human	Forward:	CACCGACTGCTGGAACTTCGCTCAC	lentiCRISPR
		Reverse:	AAACGTGAGCGAAGTTCCAGCAGTC	
APC	Human	Forward:	CCTGTGAGCGAAGTTCCAGCAGTGTCCGAT	pARv-RFP
		Reverse:	CGGACACTGCTGGAACTTCGCTCACAGG	
MSX1	Human	Forward:	CACCGAGGCGCTCATGGCCGACCAC	lentiCRISPR#1
		Reverse:	AAACGTGGTCGGCCATGAGCGCCTC	
MSX1	Human	Forward:	CACCGCCCACCGAGAAATGGCCGAG	lentiCRISPR#2
		Reverse:	AAACCTCGGCCATTTCTCGGTGGGC	
MSX1	Human	Forward:	TGGAGGCGCTCATGGCCGACCACAGGCGAT	pARv-RFP
		Reverse:	CGCCTGTGGTCGGCCATGAGCGCCTCCA	

List of primers used for analysis of gene regulatory regions occupancy by MSX1

Gene	Organism	Primer sequence (5´ to 3´)
ASCL2	Human	Forward: GCCTGCTTTTGTATTGCCCA
		Reverse: AGTTTCAGCCTCCCGAGTAG
ASCL2	Human	Forward: GACGGCTCAGATAGTGTGGA
		Reverse: CACCACCAACACCTCTCTCT
SP5	Human	Forward: CCCCCTTTGATCAGGAAAAT
		Reverse: GCTTCAGGATCACCTCCAAG
SOX17	Human	Forward: AGCTTCTTGGTGCGCTAGTC
		Reverse: GGGGGAAACAACTTTCACAA

List of primers used for amplification of CRISPR-targeted sites from genomic DNA

Gene	Organism	Primer sequence (5´ to 3´)
MSX1	Human	Forward: CACTACAGGAAGCTAGCTTCTTCCCGCAAGG
		Reverse: GGCAAAGAAGTCATGTCAGCAGCCGGGGCC
MSX1	Human	Forward: GGCTGGCCAGTGCTGCGGCAGAAGGG
		Reverse: CACGCCATTGAAATCTGGCTGCTATTATGCCGAG

List of primers used for amplification of *SP5* promoter element

Gene	Organism	Primer sequence (5´ to 3´)
SP5	Human	Forward: GCGGGTACCGCGAGGGTGCAGGGTGTGCAAGTAAA
		Reverse: GCACGGAGTACCAGGAGAGA

List of primers used for amplification and mutation of *Sox17* promoter element

Gene	Organism	Primer sequence (5´ to 3´)
Sox17	Mouse	Forward: TCAAGATCTGAGGTAAAGTCCAGTCCTAAG
		Reverse: TCTAAGCTTAGGCAAATTCTAATTCATCTG
Sox17	Mouse	Forward: GGTAAATTCCCTCCTCTCTCCGTGGTCCAG
		Reverse: AGAGAGGAGGGAATTTACCTCCGTGTTACC

Table 3 | List of primary antibodies used in the study.

Name	Source	Clonality	No.	Manufacturer
α-tubulin	rabbit	polyclonal		kindly provided by L. Andera
APC	mouse	monoclonal, clone FE9	OP44	Calbiochem
β-actin	rabbit	polyclonal	A2066	Sigma-Aldrich
β-catenin	rabbit	monoclonal	#8480	Cell Signaling
GFP	mouse	monoclonal	JL-8 632381	Clontech
GFP	rabbit	polyclonal	ab290	Abcam
H3K27me3	mouse	monoclonal	#9733	Cell Signaling
Lysozyme	rabbit	polyclonal	A0099	Dako
MSX1	goat	polyclonal	AF5045	R&D Systems
Mucin 2	rabbit	polyclonal	sc-15334	Santa Cruz
PCNA	rabbit	polyclonal	ab18197	Abcam
Sox17	rabbit	polyclonal	ab83258	Abcam
Tcf7l2	rabbit	monoclonal	#2569	Cell Signaling

4 Results

The results are divided into four main chapters. The first chapter deals with Msx1 responsiveness to Wnt/β-catenin signaling in the mouse intestinal epithelium and human cell lines. In the second part, the Msx1 function in the mouse intestines and its role in tumor formation and morphology is described and an Msx1-dependent gene signature is suggested. The third chapter is dedicated to MSX1 function in human colorectal cancer and searching for MSX1-regulated genes. Finally, the last chapter describes function of the Hic1 tumor suppressor in mouse intestines.

4.1 Msx1 expression is activated by aberrant Wnt/β-catenin signaling

In order to identify genes affected by Apc loss in mouse intestinal epithelial cells, we performed gene expression profiling of the small intestinal and colonic crypts isolated from *Apc*$^{cKO/cKO}$ *Villin-CreERT2* mice. The *Apc*$^{cKO/cKO}$ mice harbor conditional knock-out alleles of the *Apc* gene with exon 14 flanked by two *loxP* sites. Cre-mediated recombination results in the excision of the exon, which shifts the reading frame downstream from the deletion. Subsequently, truncated and nonfunctional Apc protein is synthesized[182]. The *Villin-CreERT2* mice express the regulated Cre recombinase from the murine *villin* gene promoter active in all intestinal epithelial cells. Consequently, tamoxifen-induced Cre activation in *Apc*$^{cKO/cKO}$ *Villin-CreERT2* mice results in Apc depletion in the entire epithelium[62]. As early as two days after Apc loss, massive expansion of the crypt compartment was visible in the small intestine; the colonic epithelium was affected to a much lesser extent (Figure 15A). To analyze the gene expression profile, RNA samples obtained from freshly isolated crypts prior to and at day 2 and 4 after Apc loss were analyzed by DNA microarray hybridization. Increased expression of a Wnt target gene and ISC marker tumor necrosis factor receptor superfamily, member 19 (*Tnfrsf19* alias *Troy*) was detected in the small intestine already at day 2. Robust upregulation of other Wnt-responsive genes *Lgr5*, *Ascl2*, *Axin2*, and *SP5* was observed in both the small intestine and the colon at day 4. In accordance with previously published data, Paneth-cell specific markers lysozyme 1 (*Lyz1*) and defensins (*Defa6*, *Defa26*) were elevated in the colon 2 days after Apc inactivation[3]. The gene encoding transcription factor msh homeobox 1 (*Msx1*) displayed robust upregulation in the small intestine 4 days upon Apc inactivation; the change in expression was lower in the colon [the binary logarithm of fold change (logFC) 0.77 vs. 3.53; Figure 15B]. A list of twenty-

five small intestinal genes with the most changed (increased or decreased) expression (the signicifance criteria: $|\log FC| \geq 1$ and q-value < 0.05) 2 and 4 days after Apc inactivation are given in Table 4 and Table 5, respectively. Analogously, differentially expressed genes in the colon are listed in Table 6 and Table 7; for a complete list of differentially expressed genes, see reference[114]. Results of the gene expression profiling were confirmed by quantitative real-time polymerase chain reaction (qRT-PCR) analysis of mRNA produced in the small intestinal and colonic crypts (Figure 15C, left and middle). Moreover, qRT-

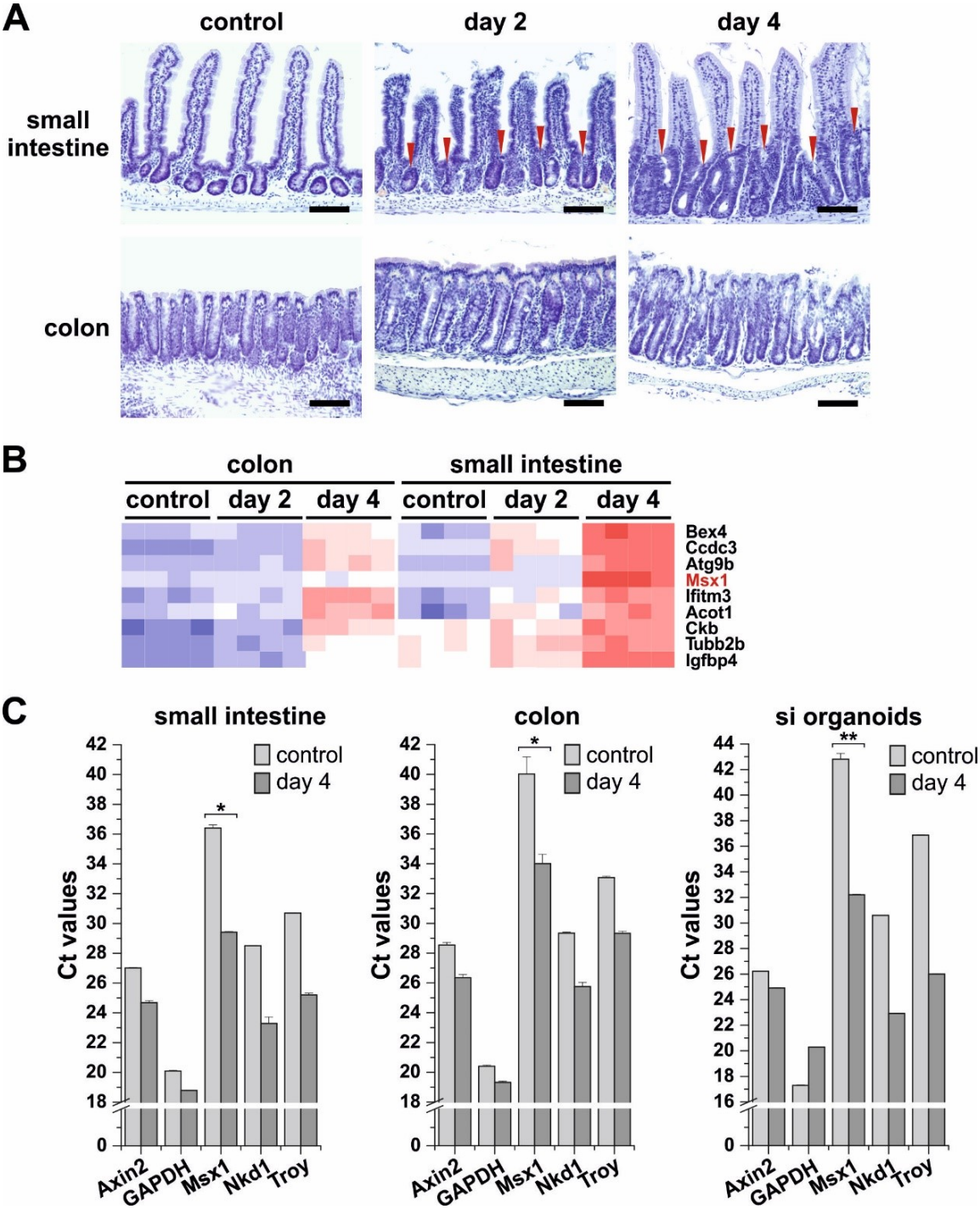


Figure 15 | *Msx1* expression increases upon Apc inactivation in the mouse intestinal epithelium. (A) Crypt hyperplasia in the *Apc*$^{cKO/cKO}$ *Villin-CreERT2* small intestine 2 and 4 days after tamoxifen administration; control tissues were obtained prior to tamoxifen treatment. Paraffin sections stained with hematoxylin (blue nuclear signal) at indicated timepoints after tamoxifen treatment are shown. Red arrowheads indicate hyperproliferative crypts; scale bar: 0.15 mm. (B) Expression profiling of crypt cells isolated from *Apc*$^{cKO/cKO}$ *Villin-CreERT2* mice 2 and 4 days after tamoxifen administration; control RNA was isolated from crypt cells with intact *Apc*. A part of the heatmap is shown, displaying robust upregulation of the *Msx1* gene after Apc inactivation in the small intestinal and colonic crypts. Subset of genes meeting significance criteria $|\log FC| \geq 1$ and q-value < 0.05 and exhibiting the highest or lowest expression after Apc inactivation is listed in Table 4 and Table 5 (small intestine) and in Table 6 and Table 7 (colon). A complete list of differentially expressed genes may be found in the Supplementary materials attached to the reference[114]. (C) Quantitative RT-PCR analysis of crypt cells isolated from *Apc*$^{cKO/cKO}$ *Villin-CreERT2* mice 4 days after tamoxifen administration confirms a significant increase in the *Msx1* mRNA expression levels in both tissues after *Apc* inactivation; control RNA was isolated from animals treated with the solvent only. Similarly, qRT-PCR analysis of *Apc*$^{cKO/cKO}$ *Villin-CreERT2* small intestinal organoids four days after 4-hydroxytamoxifen (4-OHT) treatment displays *Msx1* elevation. The diagrams show threshold cycle (Ct) values normalized to *β-actin* gene expression (Ct value of the *β-actin* gene was arbitrarily set to 17 in this and other diagrams). Wnt-responsive genes *Axin2*, *Nkd1*, and *Troy* are displayed; *GAPDH* was used as an additional housekeeping gene. RNA was obtained from four control animals and four animals treated with tamoxifen; qRT-PCR reactions were run in technical triplicates. The diagrams show representative results obtained from one animal. Error bars indicate standard deviations (SDs); *, $p < 0.05$; **, $p < 0.01$.

PCR analysis was also performed using mRNA prepared from the *Apc*$^{cKO/cKO}$ *Villin-CreERT2* small intestinal organoids treated with 4-OHT or with the solvent as a control (Figure 15C, right).

Table 4 | Differentialy expressed genes (|logFC| ≥ 1, q-value < 0.05) in the small intestinal epithelium 2 days after Apc depletion. The table shows differentially expressed genes in hyperplastic crypt cells isolated from *Apc*$^{cKO/cKO}$ *Villin-CreERT2* small intestine 2 days after tamoxifen when compared to control tissue obtained prior to tamoxifen treatment. The experiment was performed in four biological replicates.

ENTREZ	SYMBOL	GENENAME	logFC	q-value
66214	1190002H23Rik	RIKEN cDNA 1190002H23 gene	2.09	0,0044
237038	Nox1	NADPH oxidase 1	1.76	0,0071
193740	Hspa1a	heat shock protein 1A	1.61	0,045
110454	Ly6a	lymphocyte antigen 6 complex, locus A	1.51	0,012
26897	ILMN_223756	Mus musculus acyl-CoA thioesterase 1 (Acot1), mRNA.	1.49	0,0055
29820	Tnfrsf19	tumor necrosis factor receptor superfamily, member 19	1.45	0,00033
74186	Ccdc3	coiled-coil domain containing 3	1.43	2,00E-04
11459	Acta1	actin, alpha 1, skeletal muscle	1.42	0,01
11839	Areg	amphiregulin	1.4	0,012
66141	Ifitm3	interferon induced transmembrane protein 3	1.3	0,0038
406217	ILMN_231589	Mus musculus brain expressed gene 4 (Bex4), mRNA.	1.28	7,00E-04
226419	Dyrk3	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 3	1.2	0,014
406217	Bex4	brain expressed gene 4	1.2	0,028
15985	Cd79b	CD79B antigen	1.18	0,031
72821	Scn2b	sodium channel, voltage-gated, type II, beta	1.17	0,0052
26897	Acot1	acyl-CoA thioesterase 1	1.17	0,0044
73710	Tubb2b	tubulin, beta 2B	1.15	0,0017
53605	ILMN_214137	Mus musculus nucleosome assembly protein 1-like 1 (Nap1l1), mRNA.	1.14	0,0013
244886	AI118078	expressed sequence AI118078	1.14	0,0043
19703	Renbp	renin binding protein	1.12	0,021
231327	Ppat	phosphoribosyl pyrophosphate amidotransferase	1.08	0,0019
14859	Gsta3	glutathione S-transferase, alpha 3	1.08	0,018
68026	2810417H13Rik	RIKEN cDNA 2810417H13 gene	1.06	0,018
71805	Nup93	nucleoporin 93	1.05	0,0038
320685	Dctd	dCMP deaminase	1.02	0,036
67405	Nts	neurotensin	-1.33	0,01
232409	Clec2e	C-type lectin domain family 2, member e	-1.33	0,0044
20526	Slc2a2	solute carrier family 2 (facilitated glucose transporter), member 2	-1.34	0,0058
109731	Maob	monoamine oxidase B	-1.34	0,01
12013	Bach1	BTB and CNC homology 1	-1.34	0,031
228775	Trib3	tribbles homolog 3 (Drosophila)	-1.37	0,0019
13370	Dio1	deiodinase, iodothyronine, type I	-1.37	0,0042
56643	Slc15a1	solute carrier family 15 (oligopeptide transporter), member 1	-1.4	0,029
19692	Reg1	regenerating islet-derived 1	-1.41	0,0058
67082	1700011H14Rik	RIKEN cDNA 1700011H14 gene	-1.41	0,0058
70261	2010110P09Rik	RIKEN cDNA 2010110P09 gene	-1.45	0,0058
17380	ILMN_220122	Mus musculus membrane metallo endopeptidase (Mme), mRNA.	-1.5	0,012
18604	Pdk2	pyruvate dehydrogenase kinase, isoenzyme 2	-1.51	0,00019
230163	Aldob	aldolase B, fructose-bisphosphate	-1.52	0,022
64452	Slc5a4a	solute carrier family 5, member 4a	-1.54	0,0016
17380	Mme	membrane metallo endopeptidase	-1.62	0,0055
68947	Chst8	carbohydrate (N-acetylgalactosamine 4-0) sulfotransferase 8	-1.66	0,0033
54150	Rdh7	retinol dehydrogenase 7	-1.68	2,00E-04
71584	Gdpd2	glycerophosphodiester phosphodiesterase domain containing 2	-1.71	2,10E-06
13419	Dnase1	deoxyribonuclease I	-1.75	0,028
20526	Slc2a2	solute carrier family 2 (facilitated glucose transporter), member 2	-1.83	0,013
64454	Slc5a4b	solute carrier family 5 (neutral amino acid transporters, system A), member 4b	-1.89	3,50E-05
54150	Rdh7	retinol dehydrogenase 7	-1.94	2,00E-04
545156	Kalrn	kalirin, RhoGEF kinase	-1.94	0,00084
434203	ILMN_244381	Mus musculus solute carrier family 28, member 1 (Slc28a1), mRNA.	-2.14	0,00046

Table 5 | Differentialy expressed genes (|logFC| ≥1, q-value < 0.05) in the small intestinal epithelium 4 days after Apc depletion. The table shows differentially expressed genes in hyperplastic crypt cells isolated from *Apc$^{cKO/cKO}$ Villin-CreERT2* small intestine 4 days after tamoxifen administration when compared to control tissue obtained prior to tamoxifen treatment. The experiment was performed in four biological replicates.

ENTREZ	SYMBOL	GENENAME	logFC	q-value
213948	Atg9b	ATG9 autophagy related 9 homolog B (S. cerevisiae)	4,12	9,00E-08
74186	Ccdc3	coiled-coil domain containing 3	3,91	3,80E-11
66214	1190002H23Rik	RIKEN cDNA 1190002H23 gene	3,69	4,90E-07
17701	Msx1	homeobox, msh-like 1	3,53	4,10E-11
110454	Ly6a	lymphocyte antigen 6 complex, locus A	3,52	1,90E-07
16010	Igfbp4	insulin-like growth factor binding protein 4	3,48	4,80E-09
406217	ILMN_231589	Mus musculus brain expressed gene 4 (Bex4), mRNA.	3,4	4,30E-10
66141	Ifitm3	interferon induced transmembrane protein 3	3,1	1,40E-08
18612	Etv4	ets variant gene 4 (E1A enhancer binding protein, E1AF)	3,1	8,90E-07
16010	Igfbp4	insulin-like growth factor binding protein 4	3,09	4,40E-07
237038	Nox1	NADPH oxidase 1	3,06	1,80E-06
26897	ILMN_223756	Mus musculus acyl-CoA thioesterase 1 (Acot1), mRNA.	3,06	1,50E-07
29820	Tnfrsf19	tumor necrosis factor receptor superfamily, member 19	2,95	2,10E-09
406217	Bex4	brain expressed gene 4	2,81	1,20E-06
73710	Tubb2b	tubulin, beta 2B	2,66	6,30E-09
18383	Tnfrsf11b	tumor necrosis factor receptor superfamily, member 11b (osteoprotegerin)	2,63	6,90E-08
26897	Acot1	acyl-CoA thioesterase 1	2,61	4,30E-08
94179	Krt23	keratin 23	2,53	3,50E-07
14859	Gsta3	glutathione S-transferase, alpha 3	2,51	4,90E-07
103551	E130012A19Rik	RIKEN cDNA E130012A19 gene	2,5	1,00E-06
12505	Cd44	CD44 antigen	2,49	4,10E-08
27279	Tnfrsf12a	tumor necrosis factor receptor superfamily, member 12a	2,47	2,50E-05
16918	Mycl1	v-myc myelocytomatosis viral oncogene homolog 1, lung carcinoma derived	2,39	3,10E-07
13401	Dmwd	dystrophia myotonica-containing WD repeat motif	2,37	5,00E-09
109731	Maob	monoamine oxidase B	-3,76	2,30E-08
20259	Scin	scinderin	-3,77	2,50E-08
433470	AA467197	expressed sequence AA467197	-3,78	4,80E-08
216225	Slc5a8	solute carrier family 5 (iodide transporter), member 8	-3,83	9,00E-08
331063	Gsdmc2	gasdermin C2	-3,84	5,70E-07
53315	Sult1d1	sulfotransferase family 1D, member 1	-3,87	1,30E-08
11997	Akr1b7	aldo-keto reductase family 1, member B7	-3,87	1,40E-07
68396	Nat8	N-acetyltransferase 8 (GCN5-related, putative)	-3,97	3,40E-08
54447	Asah2	N-acylsphingosine amidohydrolase 2	-4,03	4,10E-08
237636	Npc1l1	NPC1-like 1	-4,12	2,60E-07
238011	Enpp7	ectonucleotide pyrophosphatase/phosphodiesterase 7	-4,14	2,50E-08
64454	Slc5a4b	solute carrier family 5 (neutral amino acid transporters, system A), member 4b	-4,18	4,20E-11
331063	ILMN_196357	Mus musculus expressed sequence AI987692 (AI987692), mRNA.	-4,22	1,40E-06
102294	Cyp4v3	cytochrome P450, family 4, subfamily v, polypeptide 3	-4,3	1,80E-06
107375	Slc25a45	solute carrier family 25, member 45	-4,36	1,00E-07
11847	Arg2	arginase type II	-4,37	7,00E-08
20363	Sepp1	selenoprotein P, plasma, 1	-4,52	1,10E-06
59020	Pdzk1	PDZ domain containing 1	-4,52	5,30E-05
20363	Sepp1	selenoprotein P, plasma, 1	-4,58	5,40E-07
238011	Enpp7	ectonucleotide pyrophosphatase/phosphodiesterase 7	-4,58	3,40E-08
21810	Tgfbi	transforming growth factor, beta induced	-4,63	2,30E-07
69983	Sis	sucrase isomaltase (alpha-glucosidase)	-4,63	8,20E-06
11814	Apoc3	apolipoprotein C-III	-5,08	4,40E-05
100045250	ILMN_196357	PREDICTED: Mus musculus hypothetical protein LOC100045250, misc RNA.	-5,14	4,60E-07
13419	Dnase1	deoxyribonuclease I	-5,97	2,30E-08

Table 6 | Differentialy expressed genes (|logFC| ≥ 1, q-value < 0.05) in the colonic epithelium 2 days after Apc depletion. The table shows differentially expressed genes in the hyperplastic colonic epithelium isolated from *Apc$^{cKO/cKO}$ Villin-CreERT2* mice 2 days after tamoxifen administration; control tissues were obtained prior to tamoxifen treatment. The experiment was performed in four biological replicates.

ENTREZ	SYMBOL	GENENAME	logFC	q-value
626708	Defa26	defensin, alpha, 26	2.38	5,00E-04
13218	Defa-rs1	defensin, alpha, related sequence 1	2.34	0,028
68009	ILMN_196346	Mus musculus defensin related cryptdin 20 (Defcr20), mRNA.	2.28	0,0037
13239	ILMN_196558	Mus musculus defensin related cryptdin 5 (Defcr5), mRNA.	2.11	0,026
68009	ILMN_196346	Mus musculus defensin related cryptdin 20 (Defcr20), mRNA.	2.01	0,0037
100044291	ILMN_221210	PREDICTED: Mus musculus hypothetical protein LOC100044291, mRNA.	1.99	0,016
17110	Lyz1	lysozyme 1	1.88	0,0096
13216	ILMN_196581	Mus musculus defensin, alpha 1 (Defa1), mRNA.	1.86	0,016
13240	Defa6	defensin, alpha, 6	1.82	0,031
17110	Lyz1	lysozyme 1	1.63	0,028
17748	Mt1	metallothionein 1	1.35	0,031
11551	ILMN_190996	Mus musculus adrenergic receptor, alpha 2a (Adra2a), mRNA.	1.28	0,026
23945	Mgll	monoglyceride lipase	1.25	0,031
213391	Rassf4	Ras association (RalGDS/AF-6) domain family member 4	1.02	0,031
12231	Btn1a1	butyrophilin, subfamily 1, member A1	-1.07	0,016
16987	ILMN_187484	Mus musculus lanosterol synthase (Lss), mRNA.	-1.12	0,016
64177	Trpv6	transient receptor potential cation channel, subfamily V, member 6	-1.33	0,04

Table 7 | Differentialy expressed genes (|logFC| ≥ 1, q-value < 0.05) in the colonic epithelium 4 days after Apc depletion. The table shows differentially expressed genes in the hyperplastic colonic epithelium isolated from *Apc$^{cKO/cKO}$ Villin-CreERT2* mice 4 days after tamoxifen administration; control tissues were obtained prior to tamoxifen treatment. The experiment was performed in four biological replicates.

ENTREZ	SYMBOL	GENENAME	logFC	q-value
12709,00	ILMN_193661	Mus musculus creatine kinase, brain (Ckb), mRNA.	3,86	2,50E-08
74186,00	Ccdc3	coiled-coil domain containing 3	3,17	6,40E-10
20568,00	Slpi	secretory leukocyte peptidase inhibitor	3,09	0,00018
73710,00	Tubb2b	tubulin, beta 2B	3,03	2,50E-09
17329,00	Cxcl9	chemokine (C-X-C motif) ligand 9	2,89	3,80E-05
66141,00	Ifitm3	interferon induced transmembrane protein 3	2,86	8,10E-08
15945,00	ILMN_253583	Mus musculus chemokine (C-X-C motif) ligand 10 (Cxcl10), mRNA.	2,7	0,0024
213948,00	Atg9b	ATG9 autophagy related 9 homolog B (S. cerevisiae)	2,58	1,90E-05
14969,00	H2-Eb1	histocompatibility 2, class II antigen E beta	2,43	0,00014
14570,00	Arhgdig	Rho GDP dissociation inhibitor (GDI) gamma	2,42	3,10E-07
15930,00	Ido1	indoleamine 2,3-dioxygenase 1	2,34	0,012
15937,00	Ier3	immediate early response 3	2,22	1,70E-07
16010,00	Igfbp4	insulin-like growth factor binding protein 4	2,21	1,30E-06
26897,00	ILMN_223756	Mus musculus acyl-CoA thioesterase 1 (Acot1), mRNA.	2,21	8,00E-06
19752,00	Rnase1	ribonuclease, RNase A family, 1 (pancreatic)	2,19	2,60E-05
66214,00	1190002H23Rik	RIKEN cDNA 1190002H23 gene	2,17	0,00017
100047619,00	ILMN_219663	PREDICTED: Mus musculus similar to solute carrier family 7, member 5, misc RNA.	2,16	9,20E-07
14160,00	Lgr5	leucine rich repeat containing G protein coupled receptor 5	2,13	2,40E-07
270152,00	Amica1	adhesion molecule, interacts with CXADR antigen 1	2,12	1,40E-06
14609,00	Gja1	gap junction protein, alpha 1	1,99	5,50E-05
11459,00	Acta1	actin, alpha 1, skeletal muscle	1,99	4,90E-05
27280,00	Phlda3	pleckstrin homology-like domain, family A, member 3	1,98	0,00039
17218,00	Mcm5	minichromosome maintenance deficient 5, cell division cycle 46 (S. cerevisiae)	1,97	2,80E-08
16145,00	Igtp	interferon gamma induced GTPase	1,97	0,025
320685,00	Dctd	dCMP deaminase	1,95	3,10E-05
109791,00	Clps	colipase, pancreatic	-2,37	2,50E-05
17287,00	Mep1a	meprin 1 alpha	-2,41	0,00036
101488,00	Slco2b1	solute carrier organic anion transporter family, member 2b1	-2,52	7,30E-05
109791,00	Clps	colipase, pancreatic	-2,57	4,60E-05
21818,00	Tgm3	transglutaminase 3, E polypeptide	-2,58	4,20E-05
69083,00	Sult1c2	sulfotransferase family, cytosolic, 1C, member 2	-2,58	9,90E-08
67971,00	Tppp3	tubulin polymerization-promoting protein family member 3	-2,59	4,80E-07
53315,00	Sult1d1	sulfotransferase family 1D, member 1	-2,63	1,60E-06
56185,00	Hao2	hydroxyacid oxidase 2	-2,67	7,10E-06
20887,00	Sult1a1	sulfotransferase family 1A, phenol-preferring, member 1	-2,7	2,00E-05
393082,00	ILMN_243966	Mus musculus methyltransferase like 7A2 (Mettl7a2), mRNA.	-2,73	1,30E-09
545288,00	Cyp2c67	cytochrome P450, family 2, subfamily c, polypeptide 67	-2,77	4,40E-07
13615,00	Edn2	endothelin 2	-2,98	2,80E-05
22635,00	Zan	zonadhesin	-2,99	3,60E-08
233038,00	Nccrp1	non-specific cytotoxic cell receptor protein 1 homolog (zebrafish)	-3,11	2,10E-06
216225,00	Slc5a8	solute carrier family 5 (iodide transporter), member 8	-3,13	1,60E-06
219033,00	Ang4	angiogenin, ribonuclease A family, member 4	-3,25	6,70E-05
13107,00	Cyp2f2	cytochrome P450, family 2, subfamily f, polypeptide 2	-3,49	3,60E-08
18947,00	Pnliprp2	pancreatic lipase-related protein 2	-3,52	1,10E-05
232889,00	Pla2g4c	phospholipase A2, group IVC (cytosolic, calcium-independent)	-3,72	3,00E-04
331063,00	ILMN_196357	Mus musculus expressed sequence AI987692 (AI987692), mRNA.	-4,02	4,50E-06
18947,00	Pnliprp2	pancreatic lipase-related protein 2	-4,18	2,40E-06
270328,00	ILMN_196360	Mus musculus gasdermin C3 (Gsdmc3), mRNA.	-4,29	1,90E-07
331063,00	Gsdmc2	gasdermin C2	-4,4	3,50E-07
100045250,00	ILMN_196357	PREDICTED: Mus musculus hypothetical protein LOC100045250, misc RNA.	-4,5	3,70E-06

MSX1 gene sensitivity to activation (or inhibition) of the Wnt pathway was tested in cultured human cells. First, in human embryonic kidney (HEK293) cells, the Wnt signaling pathway was activated by Wnt3a ligand or GSK3β inhibitor (2'Z,3'E)-6-bromoindirubin-3′-oxime (BIO). Quantitative RT-PCR analysis revealed increased expression of Wnt target genes including *MSX1* and its paralog *MSX2*. However, when compared to other tested Wnt target genes, *MSX1* (and *MSX2*) expression was increased only moderately (in BIO-treated cells) or not affected at all (in Wnt3a-stimulated cells; Figure 16).

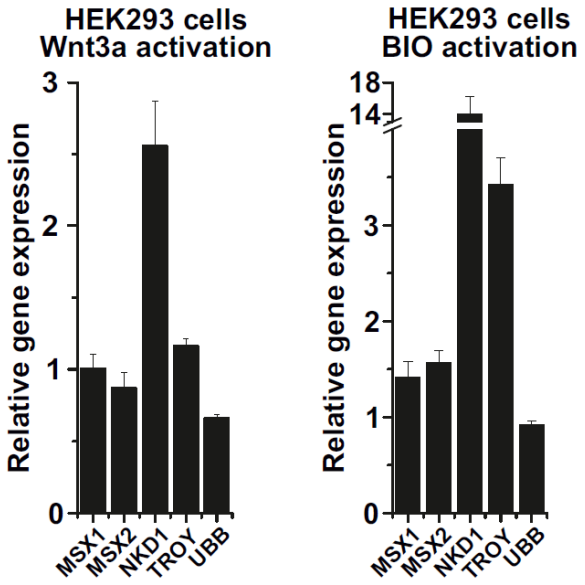


Figure 16. *MSX1* gene responsiveness to the Wnt pathway activation in HEK293 cells. Quantitative RT-PCR analysis of the *MSX1* expression level in HEK293 cells upon activation by Wnt3a (left) or GSK3β inhibitor BIO (right). The diagrams show relative gene expression levels normalized to control cells (gene expression levels were arbitrarily set to 1). Control RNA samples were obtained from cells treated with the solvent only. PCR reactions were run in triplicates; error bars indicate SDs.

Next, the Wnt pathway was activated by disruption of the *APC* gene, correspondingly to experiments performed in Apc-deficient mice. STF cells, which are HEK293 cells harboring genome-integrated Wnt-responsive luciferase reporter "Super TOP-FLASH"[411], were used for the analysis. Two STF cell lines with different truncated forms of *APC* were employed. STF cells with mutations in exon 10 were generated previously by using transcription activator-like effector nucleases (TALENs)-mediated DNA editing[372]. Using the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) system, the mutational hotspot in exon15 of the *APC* gene[219] was targeted in STF cells, generating cells producing longer APC polypeptide (Figure 17A and B). Both cell lineages displayed elevated levels of the Wnt signaling activity and increased expression of tested Wnt target genes, including *MSX1* and *MSX2*, in comparison to parental STF cells with intact APC (Figure 17C and D). On the other hand, small interfering RNA (siRNA)-mediated depletion of β-catenin mRNA in human colorectal cancer (CRC) cells SW480 and SW620 or in STF cells producing truncated APC protein led to considerable decrease of the *MSX1* and *MSX2* mRNA levels (Figure 18).

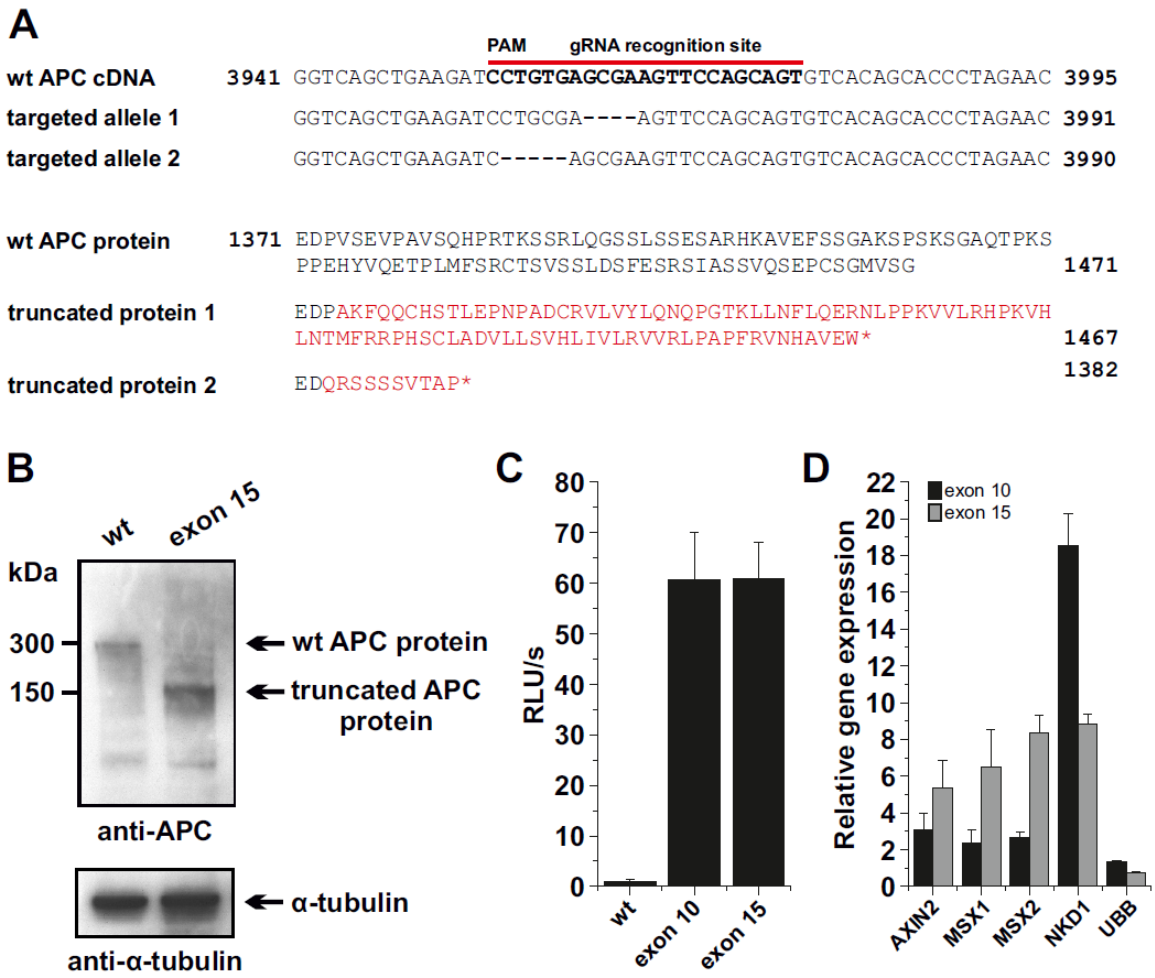


Figure 17 | STF cells producing truncated APC protein exhibit increased Wnt signaling and *MSX1* levels.

(A) DNA (top) and protein (bottom) sequences showing the CRISPR/Cas9-targeted region in the *APC* locus (exon 15). Numbers indicate the positions in the translated portion of APC cDNA or protein. The guide RNA (gRNA) recognition sequence and adjacent PAM sequence are in bold and indicated by red line. The CRISPR/Cas9-mediated cut and subsequent repair of genomic DNA by non-homologous end joining generated 4 and 5 nucleotides long deletions (depicted in "targeted allele 1" and "targeted allele 2") in the coding sequence, resulting in a premature stop of translation. Amino acid residues translated upon the frameshift are shown in red; asterisks indicate premature termini of the protein. (B) Western blotting of lysates obtained from STF cells with an anti-APC antibody confirmed production of truncated APC protein in cells harboring truncated *APC* alleles. An anti-α-tubulin antibody was used as a loading control. (C) Luciferase reporter assay in STF cells harboring truncated *APC* alleles in exon 10 and 15 reveals increased Wnt pathway activity compared to STF cells with intact APC. Relative luciferase units (RLU) indicate level of the luciferase activity normalized to cell number and viability (measured by CellTiter-Blue® Cell Viability Assay). Samples were measured in technical duplicates; experiment was performed in two replicates, representative results are shown; error bars indicate SDs. (D) Quantitative RT-PCR analysis of STF cells harboring truncated *APC* alleles in exon 10 and 15 reveals increased MSX1 expression relative to STF cells with intact APC. The diagram shows relative gene expression levels (gene expression levels in control cells were arbitrarily set to 1). PCR reactions were run in triplicates; experiment was performed in two replicates, representative results are shown; error bars indicate SDs.

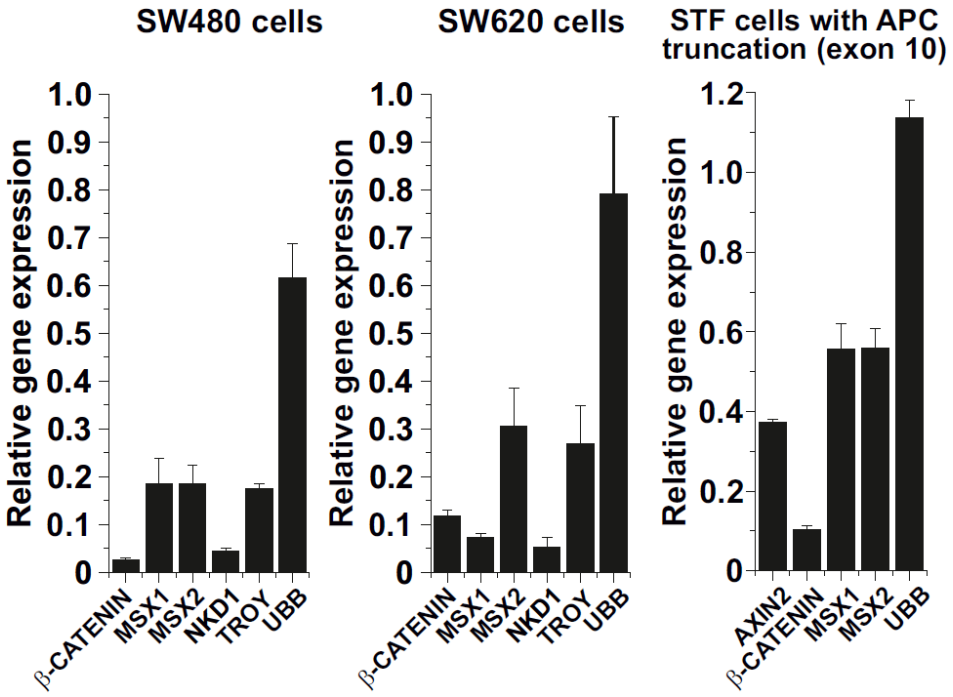


Figure 18 | *MSX1* expression increases in human cells producing truncated APC protein. Quantitative RT-PCR analysis of human colorectal cancer cells SW480 and SW620 and STF cells expressing truncated APC protein transfected with small interfering RNA (siRNA) targeting *β-catenin*. The diagrams shows gene expression levels normalized to internal housekeeping gene *β-actin* and relative to cells transfected with non-silencing siRNA (gene expression levels in control cells were arbitrarily set to 1). PCR reactions were run in triplicates; all experiments were performed in duplicates, representative results are shown; error bars indicate SDs.

4.2 The Msx1 function in the mouse intestines

To study Msx1 protein expression in the mouse intestinal epithelium, immunohistochemical staining of the small intestine and colon obtained from wild-type mice was performed using anti-Msx1 antibody. However, Msx1 protein was not detected in either tissue. Next, *Msx1* expression was tested by qRT-PCR analysis of cells isolated from *LGR5-EGFP-IRES-CreERT2* small intestinal crypts. These mice carry a knock-in allele expressing enhanced green fluorescent protein (EGFP) and regulated Cre recombinase from the promoter of intestinal stem cell marker *Lgr5*[12] (reviewed in[14]). The cells were stained with anti-CD24 (marker of cells residing at the crypt base), anti-CD45 (marker of leukocytes to exclude these cells from the analysis), and anti-EpCAM (marker of epithelial cells) antibodies. $CD45^{-}EpCAM^{+}$ cells were sorted into three populations: $CD24^{+}/GFP^{-}$, $CD24^{+}/GFP^{+}$ large cells and $CD24^{+}/GFP^{+}$ small cells (Figure 19, left). Quantitative RT-PCR analysis revealed elevated expression of stem cell markers *Lgr5* and *Troy* in $CD24^{+}/GFP^{+}$ small cells; on the other hand, $CD24^{+}/GFP^{-}$ population exhibited elevated expression of Paneth cells markers cryptidins and lysozyme. $CD24^{+}/GFP^{+}$ large "intermediate" cells retain some level of GFP protein, and therefore were gated as GFP^{+}, but don't express *Lgr5* and *Troy* at the same level as the population of $CD24^{+}/GFP^{+}$ small cells (Figure 19, right).

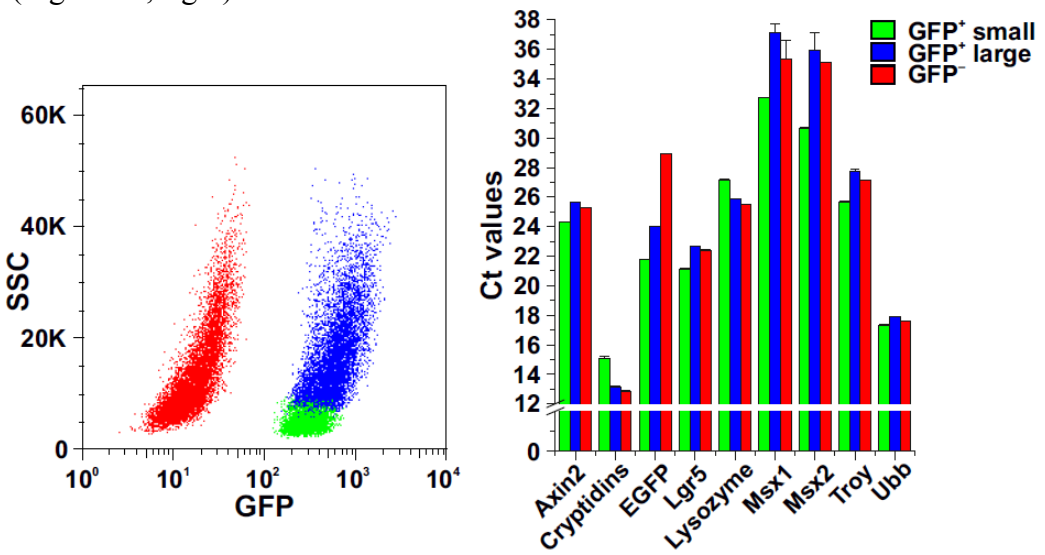


Figure 19 | Msx1 is enriched in mouse intestinal stem cells. *Lgr5-EGFP-IRES-CreERT2* small intestinal crypt cells were stained and FACS-sorted; $CD45^{-}EpCAM^{+}$ fraction was sorted into populations of $CD24^{+}/GFP^{-}$, $CD24^{+}/GFP^{+}$ large and $CD24^{+}/GFP^{+}$ small cells (left panel). Quantitative RT-PCR analysis revealed Msx1 abndance in the $CD24^{+}/GFP^{+}$ small cells (a prospective stem cell) population. Additional stem and Paneth cell marers are displayed. PCR reactions were run in triplicates; the experiment was performed three times, representative results are shown; error bars indicate SDs.

Msx1 marks ectopic crypts formed in the Apc-deficient small intestine

We further analyzed Msx1 localization in the intestinal epithelium of *Apc*$^{cKO/cKO}$ *Villin-CreERT2* mice by immunohistochemical staining at paraffin sections obtained from tissues 2, 3, 4, and 7 days upon Apc inactivation. In order to avoid premature death of the experimental mice, we lowered the tamoxifen dose to 1 mg/animal (i.e. 20 % of the amount that was used for the expression profiling). Quite unexpectedly, we observed at day 2 very rare nuclear Msx1 staining in the villi. Apparantly, these cells were not proliferating, as there was no colocalization of the Msx1 signal with the proliferating cell nuclear antigen (PCNA) staining. At day 3, Msx1-positive cells were more abundant and started to form clusters in close proximity to the crypt-villus border. Interestingly, some of the clusters contained proliferating cells. At day 4 and 7, the surface of the villi contained numerous invaginations with aberrantly proliferating cells. Of note, not all proliferating cells expressed Msx1. Interestingly, Msx1 protein specifically marked the ectopic crypts but was not detected in the hyperplastic crypt compartment (Figure 20).

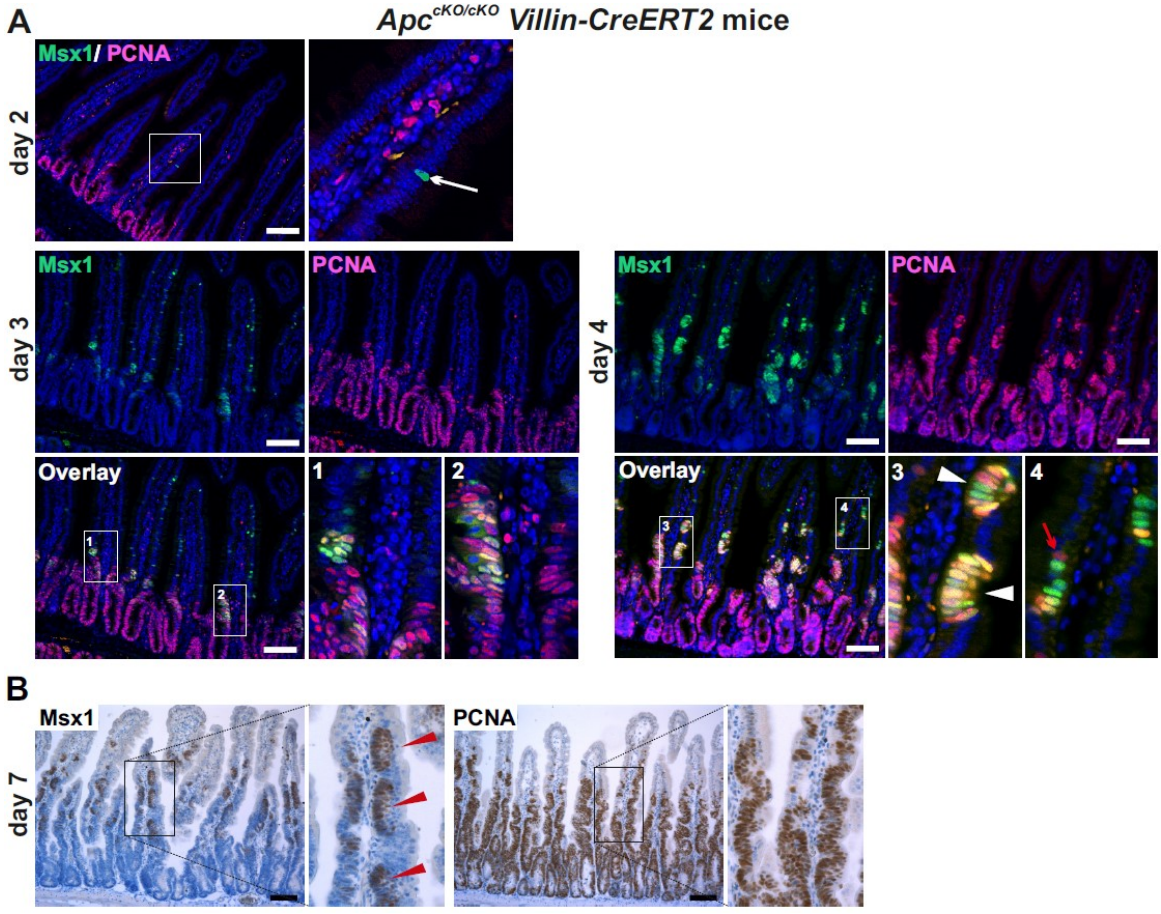


Figure 20 | *Msx1* is expressed in the ectopic crypts formed on the small intestinal villi after Apc loss. Immunohistochemical staining of Msx1 and PCNA in the *Apc*$^{cKO/cKO}$ *Villin-CreERT2* small intestine 2, 3, 4, and 7 days after tamoxifen administration. (A) Fluorescence microscopy images show Msx1 (green) and PCNA (magenta) protein localizations. At day 2, very rare Msx1-positive cells (indicated by white arrow) were detected in the villi. At day 3, Msx1-positive cell clusters with proliferating cells were observed in the villi (see enlarged sections no. 1 and 2). At day 4, pockets of proliferating (PCNA-positive) cells (i.e. ectopic crypts; indicated by white arrowhead in the enlarged section no. 3), which frequently express Msx1, are formed on villi. Some PCNA-positive cells in the villi lack Msx1 staining (indicated by red arrow in the enlarged section no. 4). The sections were counterstained with 4′,6-diamidine-2′-phenylindole dihydrochloride (DAPI; nuclear blue signal); scale bar: 0.15 mm. (B) Immunohistochemical staining reveals numerous ectopic crypts containing Msx1- and PCNA-positive cells (brown nuclear signal) at day 7. Note that the ectopic crypts have orthogonal orientation to the "normal" crypt-villus axis (indicated by red arrowheads). At least four animals were analyzed for each timepoint, representative images are shown. The sections were counterstained with hematoxylin (blue nuclear signal); scale bar = 0.3 mm.

To verify the immunohistochemical staining, a mouse strain harboring the so-called knock-out first *Msx1* allele (a reporter-tagged insertion allele), was crossed with *Apc*$^{cKO/cKO}$ *Villin-CreERT2* mice to obtain *Apc*$^{cKO/cKO}$ *Msx1*$^{+/LacZ}$ *Villin-CreERT2* animals. These mice have one "healthy" *Msx1* allele and one knock-in allele with DNA sequence encoding bacterial β-galactosidase (*lacZ*) gene downstream from the *Msx1* promoter (i. e. these mice are heretozygous for *Msx1*). The *lacZ* expression (simultaneously with *Apc* inactivation) was induced by administration of tamoxifen. Mice were sacrificed after 7 days and *lacZ* expression was visualized by X-Gal metabolite (Figure 21).

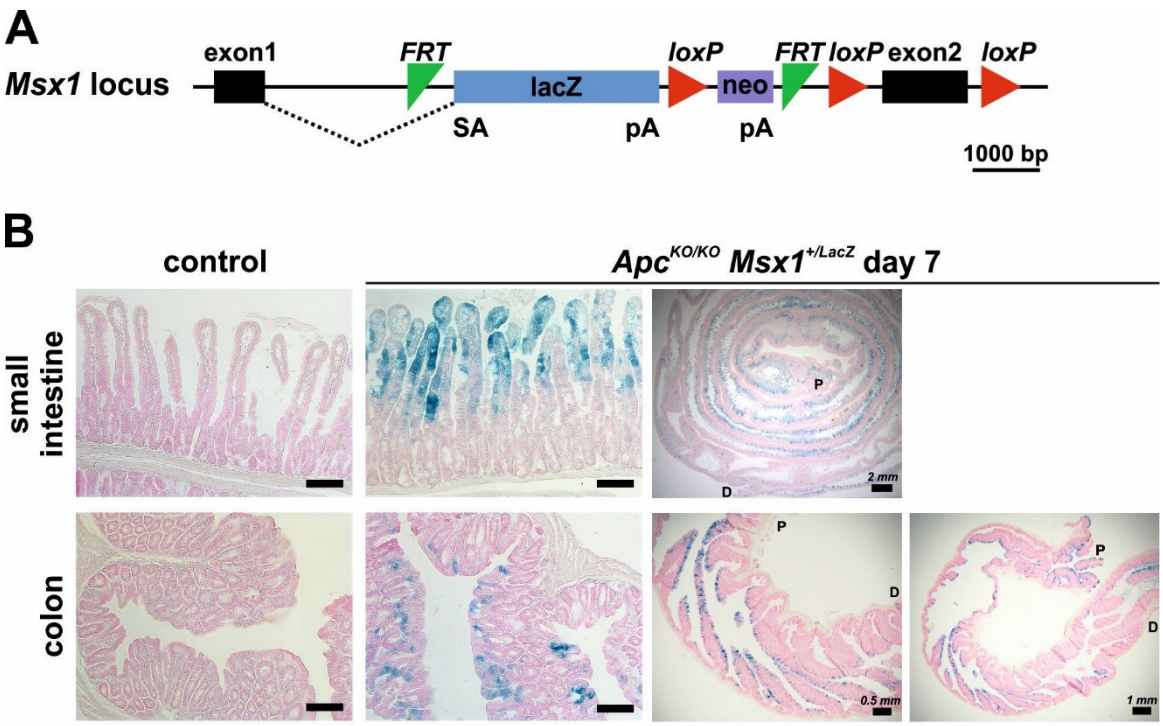


Figure 21 | Analysis of the *Msx1-LacZ* strain.
(A) The diagram shows the knock-out first *Msx1* (reporter) allele that was produced by the Knockout Mouse Program (KOMP). The *Msx1* exons are shown as black boxes, flippase recognition target (FRT) sites as green triangles, a lacZ expression reporter cassette by a blue rectangle, Cre-recognition sites (*loxP*) as red triangles, and a neomycin resistance cassette by a violet rectangle; SA, splice acceptor; pA, poly(A). The scheme was adopted from https://www.komp.org/alleles. The scalebar indicates 1000 bp length. (B) $Apc^{cKO/cKO}$ $Msx1^{+/LacZ}$ *Villin-CreERT2* mice were sacrificed 7 days after administration of a single dose of tamoxifen (1 mg per animal), control mice were administered with the solvent only. The *lacZ* expression was visualized by the β-galactosidase activity on its substrate X-gal (blue); sections were counterstained with nuclear fast red (pink). Three animals were used in the experiment, representative images are shown; scale bar: 0.15 mm (or as indicated).

To analyze the ectopic crypts in more detail, expression of two Wnt-responsive genes and intestinal stem cell markers *Olfm4* and *Ascl2* was evaluated by mRNA *in situ* hybridization. Both transcripts exhibited a predicted pattern of staining in control tissue according to previously published data[377]. *Olfm4* mRNA was detected at the base of "normal" crypts and throughout the hyperplastic crypt compartment at day 7, but not in the ectopic crypts. On the other hand, *Ascl2* mRNA was present in "normal" crypts, slightly decreased in the hyperplastic crypts, and upregulated in the ectopic crypts (Figure 22A). To investigate characteristics of the ectopic crypts, $Apc^{cKO/cKO}$ *Villin-CreERT2* mouse small intestinal epithelium was isolated 7 days after tamoxifen administration and dissociated to single cell suspension. The cells were fixed, stained with anti-epithelial cell adhesion molecule (EpCAM; marker of epithelial cells) and anti-Msx1 antibodies, and FACS sorted to obtain epithelial Msx1-positive ($Msx1^+$) and Msx1-negative ($Msx1^-$) populations (Figure 2B, left). Quantitative RT-PCR analysis revealed increased expression of Wnt target genes and stem cell markers *Ascl2*, *Axin2*, *Lgr5*, and *SP5* in $Msx1^+$ cells (Figure 22B, right). On the other hand, *Olfm4* expression was decreased in $Msx1^+$ cells, which is in compliance with mRNA staining by *in situ* hybridization (Figure 22A).

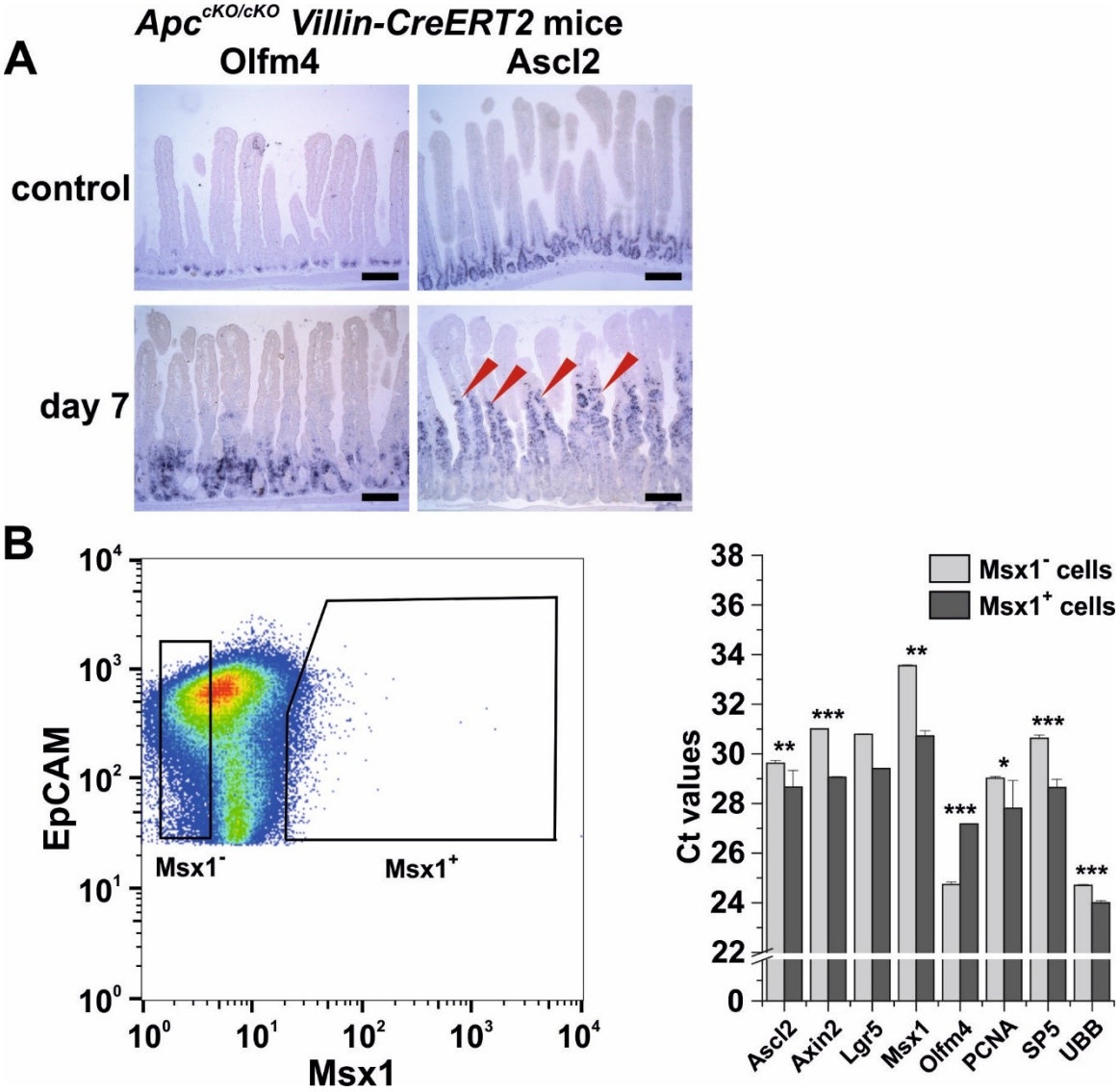


Figure 22 | Ectopic crypts express Wnt target genes and intestinal stem cell markers. (A) Detection of *Olfm4* and *Ascl*2 transcripts in the crypt hyperplasia in *Apc*$^{cKO/cKO}$ *Villin-CreERT2* small intestine 7 days after tamoxifen administration; control tissues were obtained from mice of the same genetic background prior to tamoxifen treatment. *In situ* hybridization of paraffin sections showed upregulated expression of *Olfm4* and *Ascl2* in the hyperplastic or ectopic crypts, respectively. Scale bar: 0.3 mm. (B) Single cell suspension from the small intestinal epithelium (EpCAM$^+$ cells) 7 days upon Apc depletion was stained with an anti-Msx1 antibody and an Msx1 highly positive (+) or negative (-) population, respectively, was obtained using FACS-sorting (left panel). Subsequent qRT-PCR analysis revealed upregulation of stem cell signature (*Ascl2*, *Lgr5*) and Wnt signaling-regulated genes (*Axin2*, *SP5*) in Msx1-expressing cells. The total RNA level was normalized to expression of internal housekeeping gene *β-actin* that was arbitrarily set to 23; another housekeeping gene *UBB* is shown. Error bars indicate SDs; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

Msx1 is enriched in the Apc-deficient intestinal tumors

Next, we analyzed Msx1 expression in early intestinal lesions developed in *Apc*$^{cKO/cKO}$ *Lgr5-EGFP-IRES-CreERT2* mice. These mice enable tamoxifen-induced *Apc* inactivation specifically in ISCs. Mice were sacrificed at several timepoints after tamoxifen administration and the tissues were analyzed by immunohistochemical staining. Msx1-positive cells were observed already at day 4 in proliferating enlarged small intestinal crypts (Figure 23A). At later timepoints, Msx1-positive cells were present in (micro)adenomas. Similarly to *Apc*$^{cKO/cKO}$ *Villin-CreERT2* mice, Msx1 was not detected in the "normal" crypt compartment. Moreover, not all proliferating cells were stained by the Msx1-specific

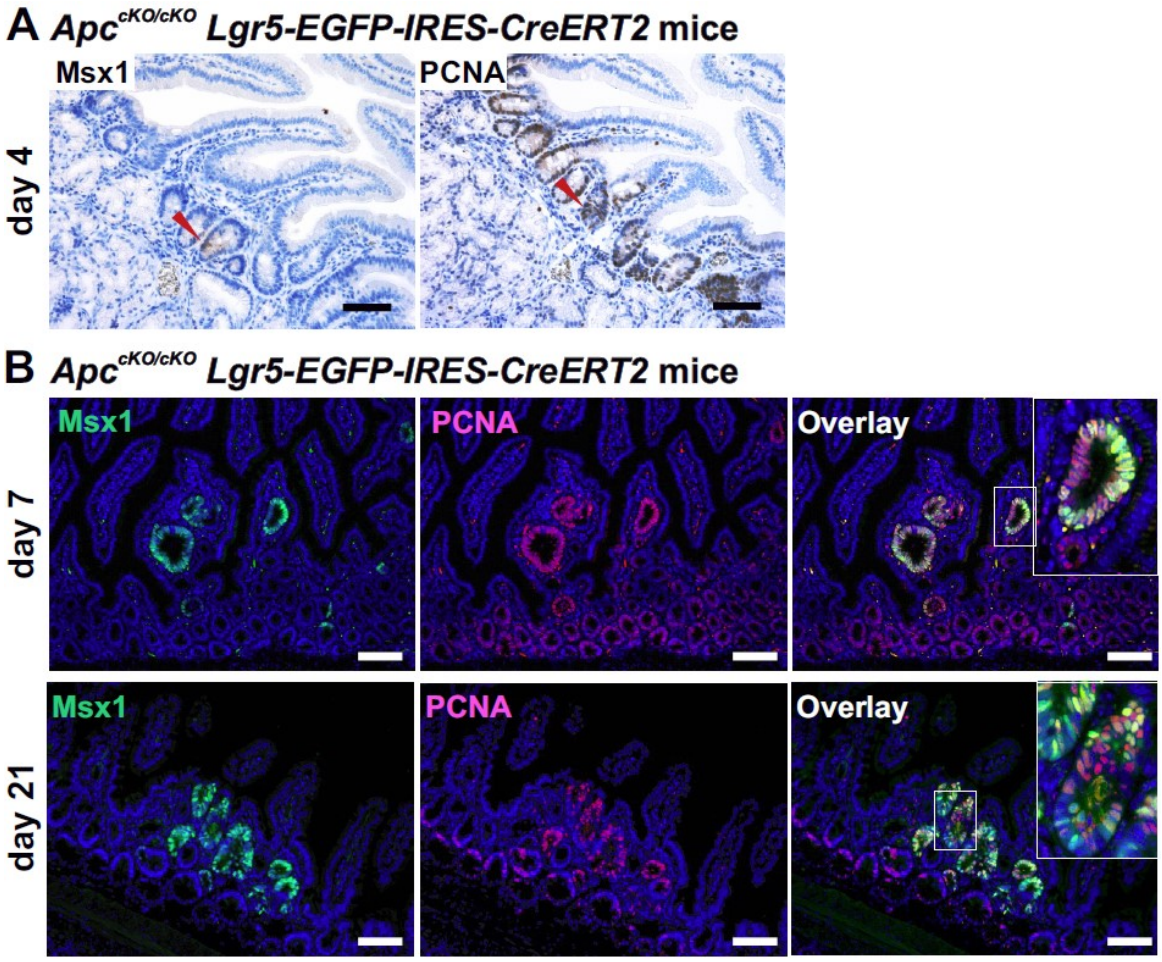


Figure 23 | Proliferating tumor cells are positive for Msx1.
Immunohistochemical staining of Msx1 and PCNA in *Apc*$^{cKO/cKO}$ *Lgr5-EGFP-IRES-CreERT2* small intestine 4, 7, and 21 days after tamoxifen administration. (A) Immunohistochemical staining reveals Msx1 and PCNA colocalization (brown nuclei) in enlarged crypts 4 days after *Apc* loss (red arrowheads). The sections were counterstained with hematoxylin; scale bar: 0.15 mm. (B) Fluorescence microscopy images show Msx1 (green) and PCNA (red) protein localizations 7 and 21 days upon Apc inactivation. The sections were counterstained with DAPI; boxed areas are magnified in the insets; scale bar: 0.3 mm.

antibody (Figure 23B). A similar pattern of *Msx1* expression was observed in intestinal tumors of *Apc*$^{+/Min}$ mice. These mice carry a nonsense mutation in one *Apc* allele and develop numerous intestinal tumors in adulthood[349]. Msx1 protein was clearly detected in upper parts of the small intestinal tumors and in colonic aberrant crypt foci (ACF), but not in the crypts (Figure 24A). The pattern of *Msx1* expression was verified by *in situ* hybridization with an *Msx1* antisense probe (Figure 24B). Finally, *Msx1* abundance in *Apc*$^{+/Min}$ intestinal tumors was confirmed by qRT-PCR analysis. Of note, the analysis of RNA isolated from multiple tumors isolated from three animals did not reveal any correlation between the level of *Msx1* expression and the tumor size or its position along the anterior-posterior axis of the small intestine (Figure 24C).

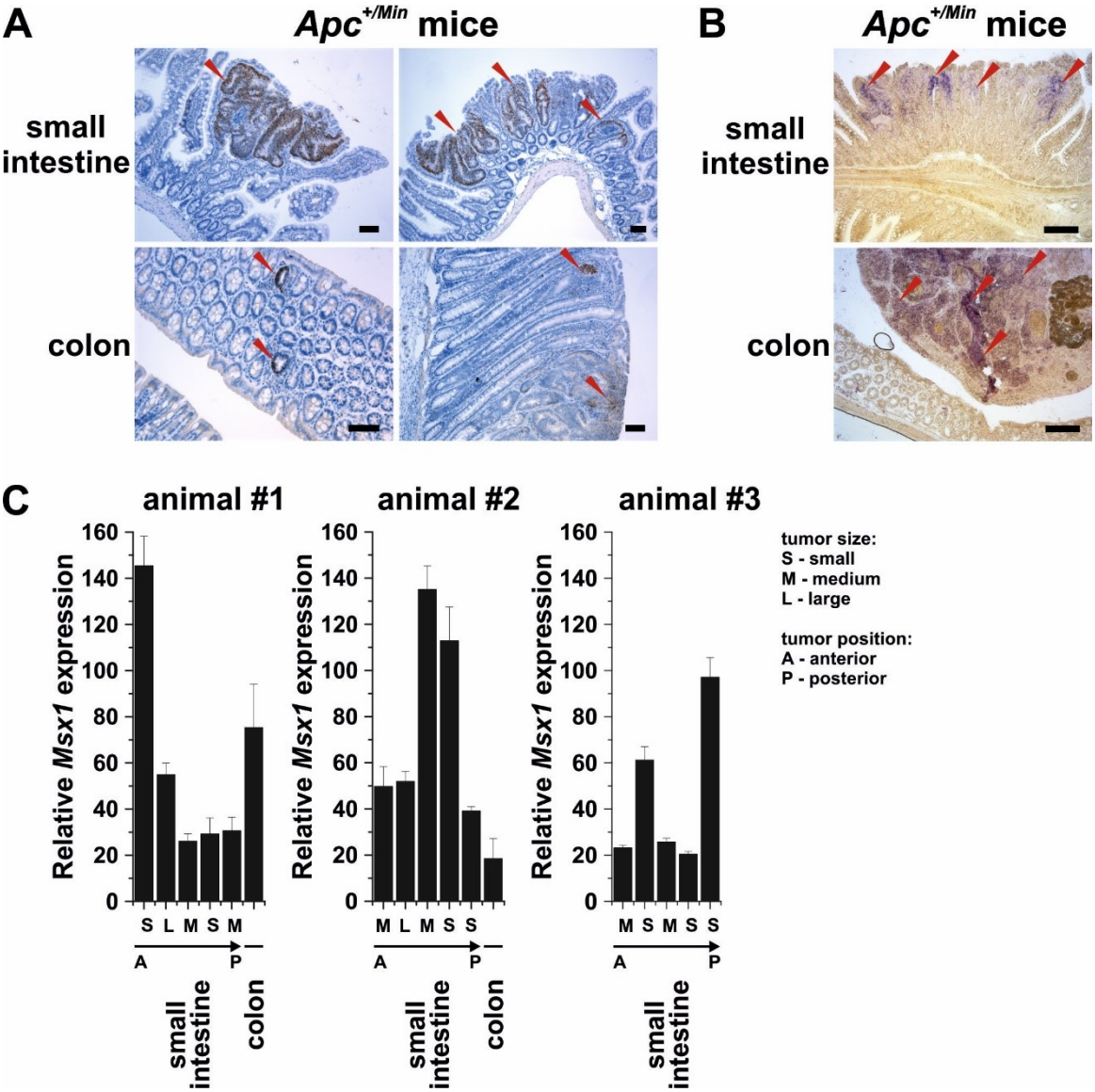


Figure 24 | Msx1 is enriched in tumors developed in *Apc*$^{+/Min}$ mouse.
(A) Immunohistochemical staining of Msx1 protein (brown nuclear signal) in two small intestinal microadenomas and colonic aberrant crypt foci (red arrowheads). The sections were counterstained with hematoxylin; scale bar: 0.15 mm. (B) *In situ* hybridization of *Msx1* mRNA (violet signal) in small intestinal and colonic tumors (red arrowheads). Scale bar: 0.3 mm. (C) Quantitative RT-PCR analysis indicates a significant increase in the *Msx1* expression levels in the mouse small intestinal and colonic tumors isolated from 20 weeks old *Apc*$^{+/Min}$ mice. The *Msx1* Ct values were normalized to *β-actin* gene expression (the *β-actin* gene Ct value was arbitrarily set to 17). The diagrams show *Msx1* expression in tumor samples relative to the control healthy mucosa (*Msx1* expression in control tissue was arbitrarily set to 1). The tumor size is indicated as S (small), M (medium), L (large); tumor position along the anterior-posterior axis of the small intestine is indicated on thy Y axis by black arrow from anterior (A) to posterior (P). The experiment was performed in technical triplicates; error bars indicate standard deviations (SDs). In animal #3, no colonic tumor was detected.

Msx1 deficiency changes morphology of the small intestinal tumors

As the whole-body knock-out of the *Msx1* gene is neonatal lethal, mice harboring conditional *Msx1* alleles (*Msx1*$^{cKO/cKO}$) were further employed. *Msx1*$^{cKO/cKO}$ mice were crossed with *Villin-Cre* mice, a strain which expresses constitutively active Cre recombinase in all epithelial cells starting at embryonic day 12.5[202]. However, no histological changes were seen in the small intestine and colon upon continuous *Msx1* inactivation and neither did Msx1 loss affect morphology or growth of intestinal organoids (Figure 25).

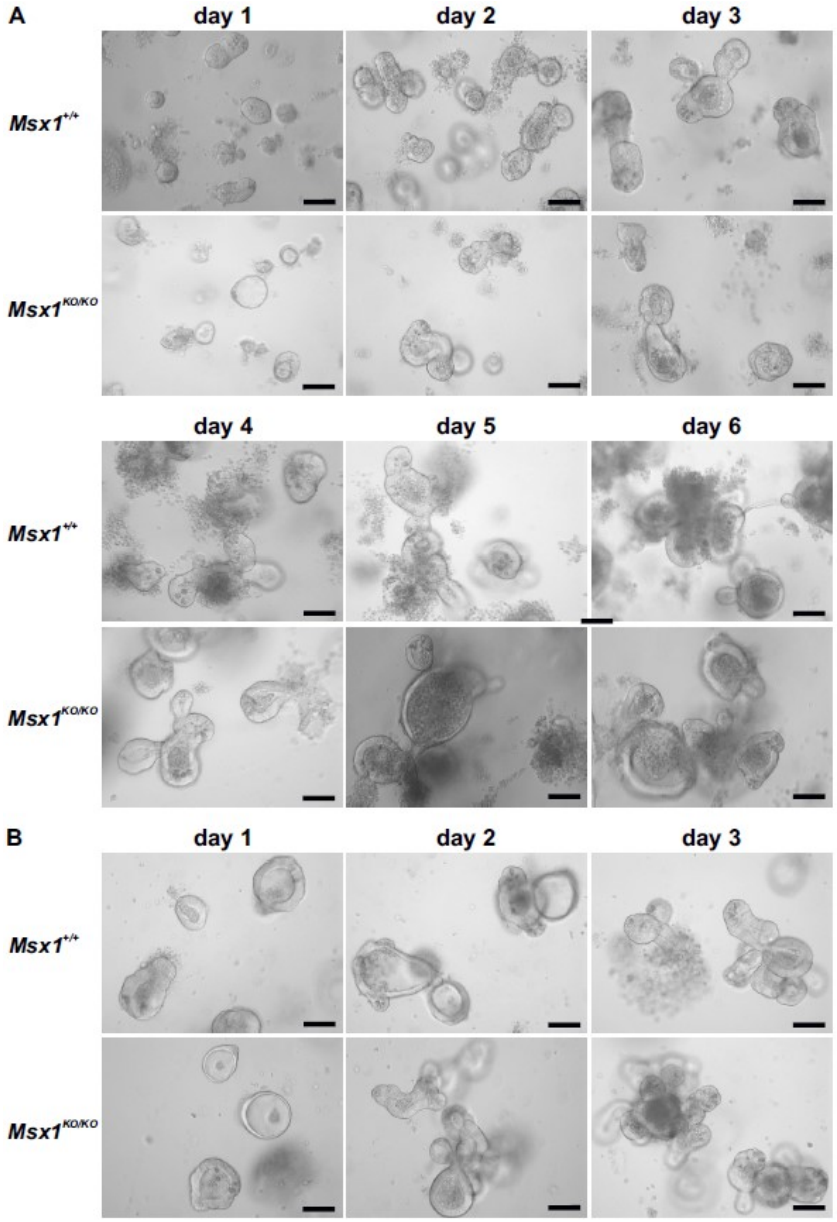


Figure 25 | Msx1 loss does not affect morphology or growth of small intestinal organoids. Stereomicroscopic images of organoids derived from the small intestine of *Msx1*$^{cKO/cKO}$ *Villin-Cre* (*Msx1*$^{KO/KO}$) and *Villin-Cre* (*Msx1*$^{+/+}$) mice. The images were taken 1, 2, 3, 4, 5, and 6 days after crypts isolation (A) and 1, 2, and 3 days after first passage (B). Scale bar: 150 µm.

In order to investigate the Msx1 function in the intestinal epithelium under non-homeostatic conditions, *Msx1*$^{cKO/cKO}$ *Villin-CreERT2* mice were generated. These mice express more efficient, tamoxifen-activated Cre recombinase and enable Msx1 inactivation throughout the entire intestinal epithelium. As mentioned above, Msx1 inactivation in homeostatic conditions does not lead to any morphological changes; therefore two models of intestinal tissue damage were applied. The *Msx1*$^{cKO/cKO}$ *Villin-CreERT2* mice were either irradiated by sublethal X-ray doses[272] to deplete proliferating cells, or administered with dextran sulfate sodium (DSS) in drinking water[121] to damage the epithelial layer. However,

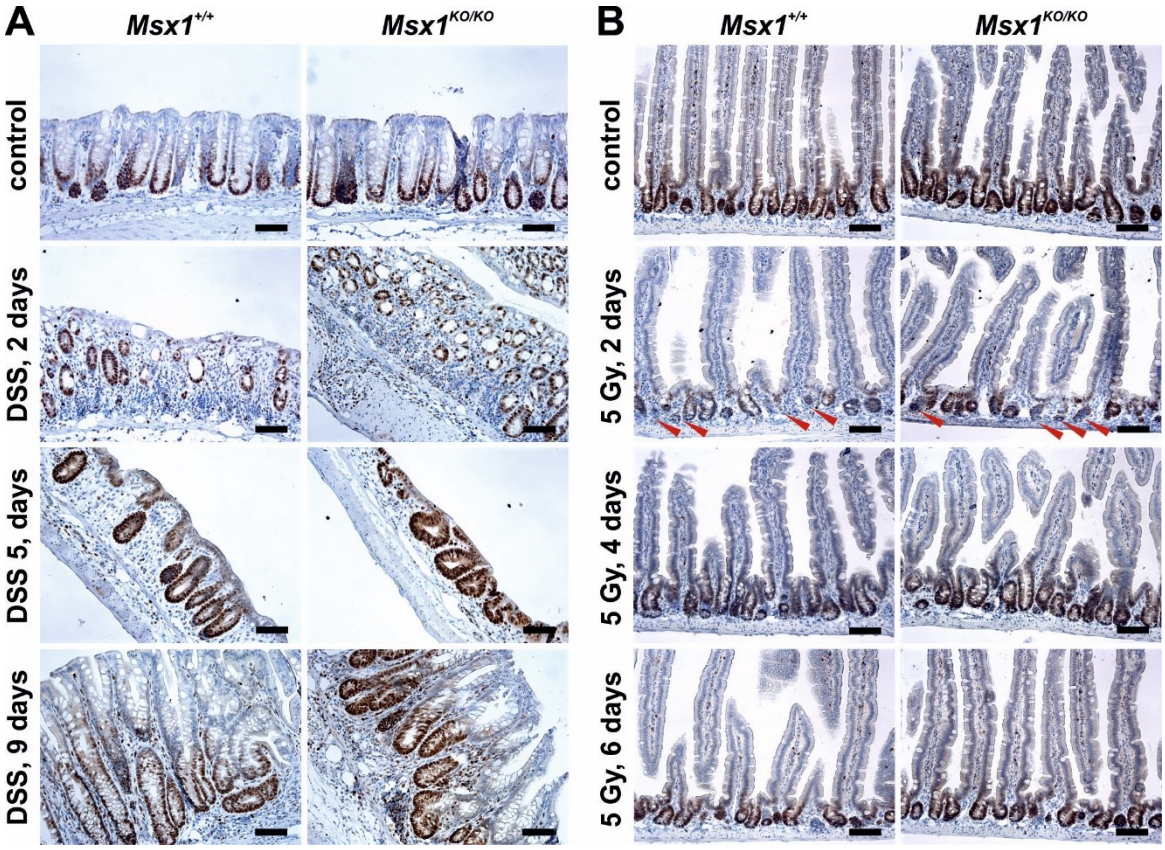


Figure 26 | Msx1 loss does not change regeneration or extent of tissue damage in the intestinal epithelium upon dextran sulfate sodium (DSS) treatment or total body irradiation. Immunohistochemical staining of PCNA in *Msx1*$^{cKO/cKO}$ *Villin-CreERT2* mice. Mice were administered with a single dose of tamoxifen (*Msx1*$^{KO/KO}$) prior to DSS treatment or irradiation to inactivate the *Msx1* gene; control animals of the same genetic bacground (*Msx1*$^{+/+}$) were administered with the solvent only. (A) DSS-induced acute colitis 2 and 5 days upon DSS withdrawal. At day 9, the colonic epithelium was completely regenerated. (B) Animals were exposed to a single dose of total body irradiation by 5 grays (Gy) and the small intestine was harvested at indicated timepoints. Of note, at day 2, number of proliferating cells in the crypts was reduced (red arrowheads) and at day 6, the epithelium was restored. Three animals of both genotypes were used for the analysis; representative images are shown. Control tissues were isolated before DSS treatment or irradiation. The sections were counterstained with hematoxylin; scale bar: 0.15 mm.

in both cases no differences regarding the epithelial regeneration or tissue damage were observed between Msx-deficient and Msx1-proficient epithelium (Figure 26).

We further analyzed mice harboring conditional alleles of both, the Msx1 and Apc genes. *Apc*$^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ *Villin-CreERT2* mice did not display any remarkable differences compared to *Apc*$^{cKO/cKO}$ *Villin-CreERT2* mice, except for the absence of *Msx1* staining, by the day 4 after tamoxifen administration; however, at day 7, Msx1 loss resulted in significant morphological change of the hyperplastic intestinal epithelium. Contrary to *Apc*$^{KO/KO}$ *Msx1*$^{+/+}$ epithelium, the PCNA-positive proliferative compartment was prolonged and reached to the top of villi (Figure 27A). Note that the ectopic crypts are not present in Msx1-deficient epithelium. Interestingly, *Ascl2* expression was detected not only in the hyperplastic or ectopic crypts, but in larger areas of the epithelial layer, often reaching to

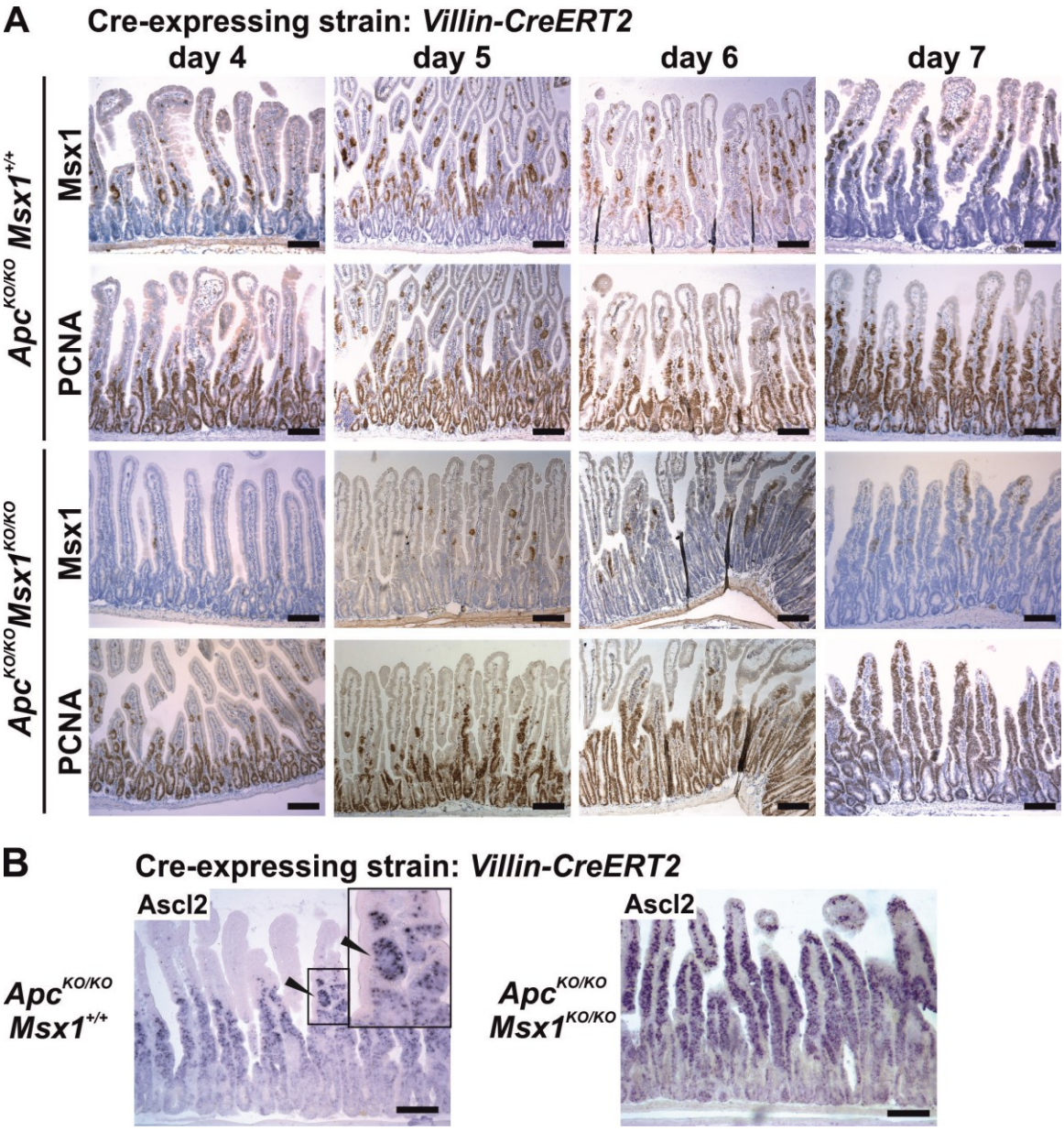


Figure 27 | Msx1 absence changes morphology of the Apc-deficient intestinal epithelium. (A) Immunohistochemical staining of Msx1 and PCNA in *Apc*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{+/+}$) and *Apc*$^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ *Villin-CreERT2 CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$) small intestine at indicated days after tamoxifen administration. As a result of Msx1 depletion, highly proliferating (PCNA-positive) cells expand from the hyperplastic crypt compartment and reach the top of the villi at day 7. Note that in Msx1-deficient epithelium the ectopic crypts are not formed and that the gene recombination is not complete (groups of Msx1-positive proliferating cells are detected on villi). The sections were counterstained with hematoxylin. (B) Detection of mRNA encoding the stem cell marker Ascl2 in *Apc*$^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$) and *Apc*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{+/+}$) intestinal epithelium using *in situ* hybridization; tissues were isolated 7 days after tamoxifen administration. While in the *Apc*$^{KO/KO}$ *Msx1*$^{+/+}$ epithelium the anti-sense *Ascl2* probe stained only the hyperplastic and ectopic crypts (black arowheads), in *Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$ mice the *Ascl2* signal robustly expands throughout the whole epithelium. Boxed area is magnified in the inset; scale bar: 0.3 mm.

the tips of the villi (Figure 27B). Increased proliferation was accompanied by loss of cell differentiation, as evidenced by the absence of trimethylation of histone H3 at lysine 27 (H3K27me3), a specific marker of differentiated epithelial cells[172] (Figure 28). Similarly, a decrease in the mRNA level and protein signal of lysozyme and mucin 2, markers of Paneth and goblet cells, respectively, was observed in the Apc-/Msx1-double-deficient epithelium in comparison to Apc-deficient epithelium (Figure 29).

Hyperproliferation of the crypt compartment upon Apc inactivation is accompanied by decrease in the percentage of cells in the G1 phase of the cell cycle. To reveal possible influence of Msx1 loss on the cell cycle progression, epithelial cells were isolated from the small intestine obtained from *Apc*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{+/+}$) and

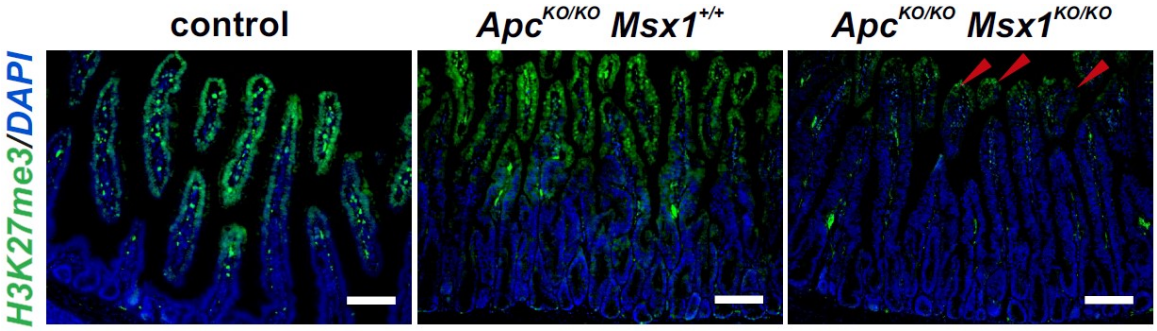


Figure 28 | Decreased trimethylation of histone H3 at lysine 27 (H3K27me3) in the Apc-/Msx1-double-deficient intestinal epithelium indicates reduced numbers of differentiated cells. Immunohistochemical staining of H3K27me3 (green signal), a marker of differentiated cells, in the *Apc*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{+/+}$) and *Apc*$^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$) small intestine 7 days after tamoxifen administration; control mice were administered with the solvent only. Note that the differentiated cells are in *Apc*$^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ *Villin-CreERT2* epithelium positioned only at the tips of the villi (red arrowheads). The sections were counterstained with DAPI; scale bar: 0.15 mm (control) and 0.3 mm (*Apc*$^{KO/KO}$ *Msx1*$^{+/+}$ and *Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$.

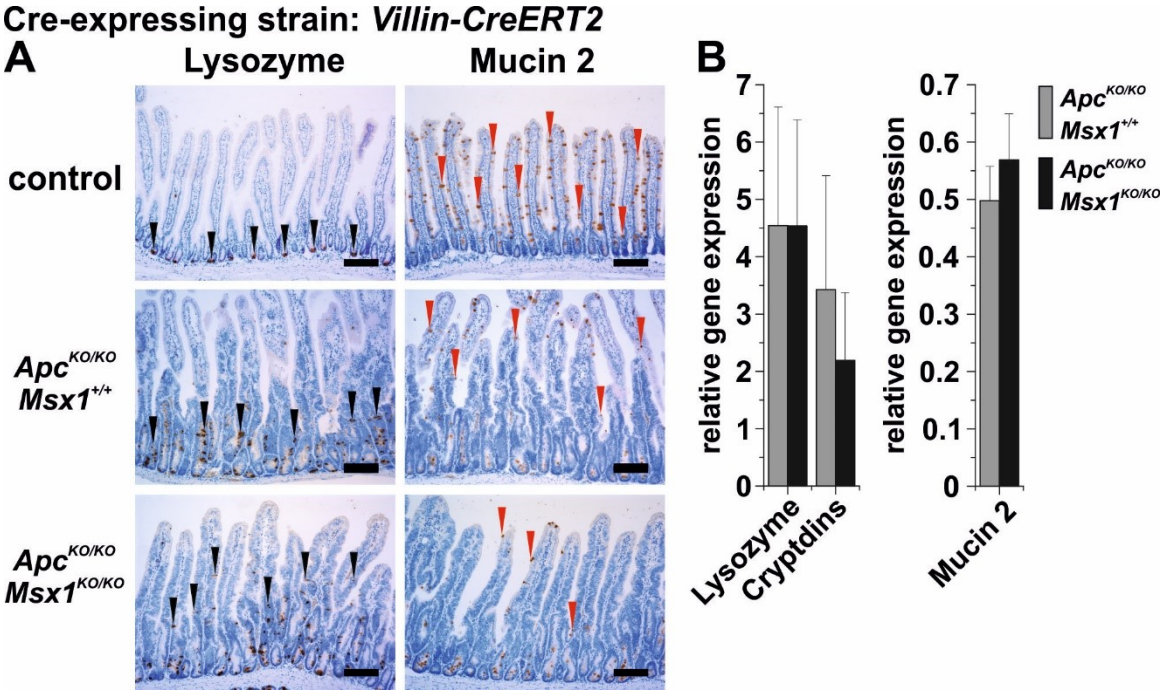


Figure 29 | Paneth and goblet cells diminish in the Msx1-deficient intestinal epithelium. *Apc*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{+/+}$) and *Apc*$^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$) mice were administered with a single dose of tamoxifen and sacrificed after 7 days; control mice were administered with the solvent only. (A) Immunohistochemical staining of lysozyme and mucin 2, markers of Paneth and goblet cells, respectively, on paraffin sections obtained from the small intestines. Note that in control epithelium, lysozyme stains only cells in the crypt base, whereas in the Apc-deficient epithelium, lysosyme-positive cells are present throughout the hyperplastic crypt compartment and in the Apc-/Msx1-double-deficient epithelium also in the villi (black arrowheads). Contrary, mucin 2-positive cells diminish in the Apc-deficient and Apc-/Msx1-double-deficient epithelium (red arrowheads). Three mice of each genotype were analysed, representative images are shown. The sections were counterstained with hematoxylin; scale bar: 0.3 mm. (B) The diagram shows expression in the *Apc*$^{KO/KO}$ *Msx1*$^{+/+}$ and *Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$ small intestine relative to wild-type tissue (the gene expression level in control mice was arbitrarily set to 1). The qRT-PCR reactions were run in technical triplicates.

Apc$^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$) mice 7 days after tamoxifen administration. The cells were fixed and stained with propidium iodide (PI) solution and the cell cycle was analyzed by flow cytometer. However, we did not observe any difference between Msx1-deficient and Msx1-proficient cells (Figure 30).

Inasmuch as the loss of Msx1 has changed the morphology of the intestinal epithelium, the effect of Msx1 loss on the life span of Apc-deficient mice was subsequently investigated. *Apc* $^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{+/+}$) and *Apc*$^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$) mice were administered with a single dose of tamoxifen and their survival was monitored. Mice with the inactivated *Apc* gene usually

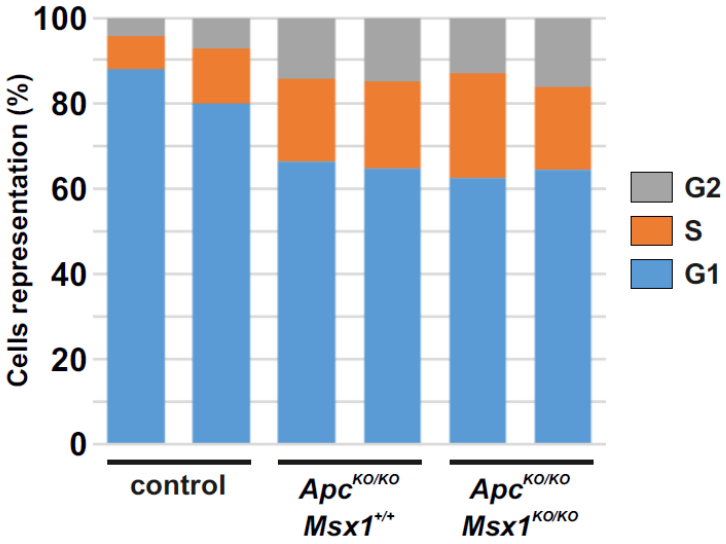


Figure 30 | Msx1 depletion does not change the percentage representation of Apc-deficient intestinal epithelial cells in individual cell cycle phases.
Apc$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{+/+}$) and *Apc*$^{cKO/cKO}$ *Msx1*$^{cKO/c}$ *Apc*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{+/+}$) and *Apc*$^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$) mice 7 days after tamoxifen administration ^{KO} *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$) mice were sacrificed 7 days after tamoxifen administration (1 mg per animal) and epithelial cells from the small intestine were isolated, fixed, and stained with propodium iodide (PI) solution. The diagram shows percentage of cells in the indicated cell cycle phases measured by flow cytometer. Two animals from each genotype were tested, control mice were administered with the solvent only. The experiment was performed twice, representative results are shown.

die within one week after tamoxifen administration due to the dysfunction of the intestinal tissue. To enable measurement over a longer period of time, the tamoxifen dose was lowered to 0.3 mg per animal, which prolonged the life span of the experimental animals. Indeed, the *Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$ mice started to die several days later than the Apc-deficient mice and the median of their survival time was higher than in *Apc*$^{KO/KO}$ *Msx1*$^{+/+}$ mice (13 and 11 days, respectively). Nevertheless, the difference in survival rates was not statistically significant (Figure 31).

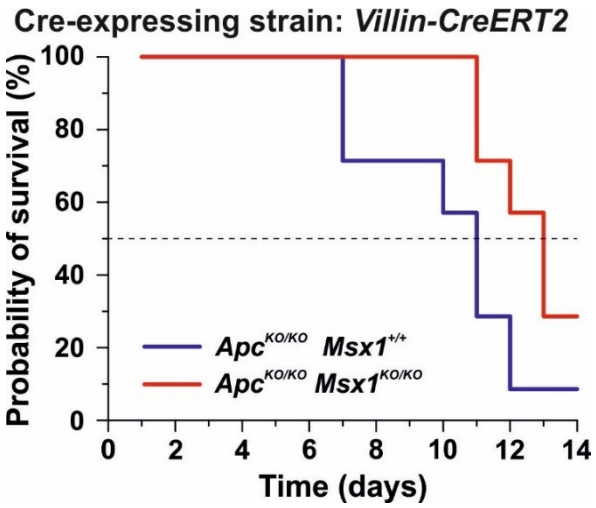


Figure 31 | Msx1 loss prolongs life span of Apc-deficient mice.
Nine *Apc* $^{cKO/cKO}$ *Villin-CreERT2* (*Apc* $^{KO/KO}$ *Msx1*$^{+/+}$) and nine *Apc* $^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$) mice were administered with a single dose of tamoxifen (0.3 mg per animal) and their survival was monitored. The diagram shows percentage of living animals at indicated timepoints.

Next, we analyzed the effect of *Msx1* gene silencing in *Apc*$^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ *LGR5-EGFP -IRES-CreERT2* mice. Msx1 loss pronouncedly changed the appearance of intestinal tumors; while the Msx1-proficient adenomas displayed a typical tubular morphology. Msx1-deficient adenomas exhibited villus-like morphology (Figure 32). As these mice survive only several weeks after tamoxifen administration, *Apc*$^{+/Min}$ *Msx1*$^{cKO/cKO}$ *Villin-Cre* mice were utilized to study more advanced tumors. To investigate differences in gene expression between Msx1-deficient and Msx1-proficient intestinal adenomas, RNA samples from tumors obtained from 20 weeks old *Apc*$^{+/Min}$ *Msx1*$^{cKO/cKO}$ *Villin-Cre* mice were analyzed. However, the differences in mRNA levels were negligible.

In order to explain the mechanism behind observed morphological changes, microarray analysis of RNA from small intestinal epithelial cells isolated from *Apc*$^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ *Villin-CreERT2* and *Apc*$^{cKO/cKO}$ *Villin-CreERT2* 7 days upon tamoxifen administration was performed. The difference in gene expression between Msx1-deficient and Msx1-proficient epithelial cells was insignificant; the significance criterium q-value < 0.05 was never reached. Nevertheless, a small set of differentially expressed genes with

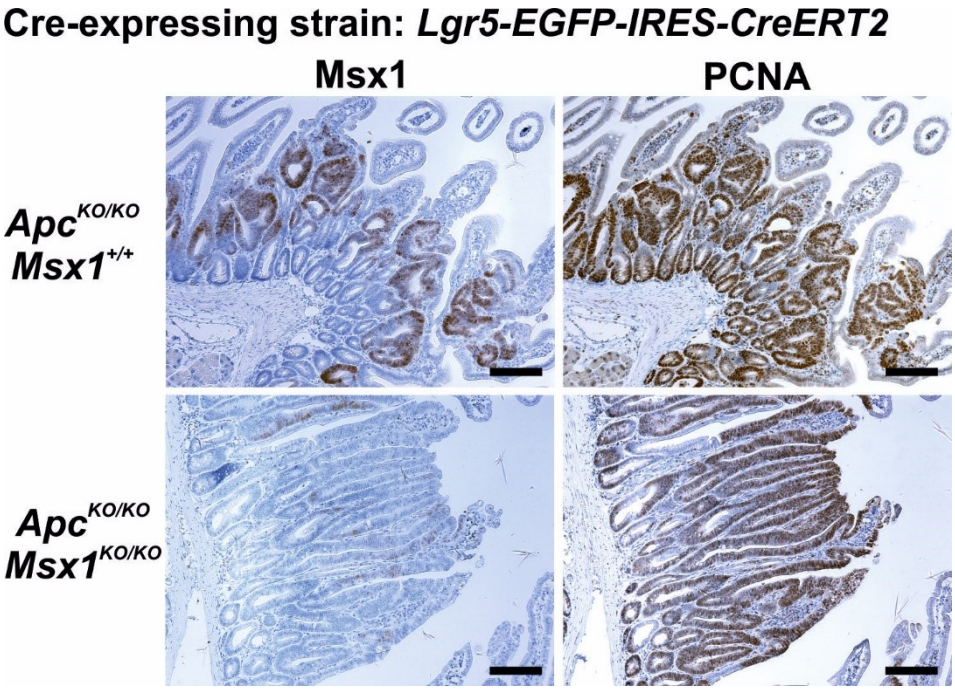


Figure 32 | Msx1 absence changes morphology of intestinal tumors. Immunohistochemical staining of Msx1 and PCNA in *Apc*$^{cKO/cKO}$ *Lgr5-EGFP-IRES-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{+/+}$) and *Apc*$^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ *Lgr5-EGFP-IRES-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$) small intestine 21 days after tamoxifen administration. Whereas the lesions in *Apc*$^{cKO/cKO}$ display the tubular phenotype, the lesions in *Apc*$^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ are transformed to predominantly villous morphology. The histological analysis of tumors was performed using samples obtained from ten animals (five for each genotype), representative images are shown. Sections were counterstained with hematoxylin; boxed area is magnified in the inset; scale bar: 0.3 mm.

significance criteria |FC| ≥ 2 and p-value < 0.05 was identified (incuding *Sox17*) and subsequently analyzed using the online tool Enrichr[123, 181] (list of twenty genes with the most increased or decreased expression is given in Table 8; for a complete list of differentially expressed genes, see reference[114]). However, the analysis did not reveal any biological process, signaling pathway, or molecular mechanism that could be involved in the observed phenotype. Quantitative RT-PCR analysis of selected Wnt target genes and markers of intestinal cell populations was performed. The analysis revealed a slight

Table 8 | Differentialy expressed genes (|logFC| ≥ 1; p ≤ 0.05) in the small intestinal Msx1 wild-type and Msx1-deficient hyperplastic epithelium.
The table shows differentially expressed genes in hyperplastic crypt cells isolated from *Apc*$^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ *Villin-CreERT2* small intestine 7 days after tamoxifen administration (1 mg per animal); control tissues were obtained prior to tamoxifen treatment. The experiment was performed in four biological replicates.

ENTREZ	SYMBOL	GENENAME	logFC	p-value
170942	Erdr1	erythroid differentiation regulator 1	3,52	2,00E-04
57742	Abhd1	abhydrolase domain containing 1	2,93	4,10E-05
57742	Abhd1	abhydrolase domain containing 1	2,84	5,60E-05
434794		Mus musculus X-linked lymphocyte-regulated 4A (Xlr4a), mRNA.	2,81	0,0083
11746	Anxa4	annexin A4	2,77	0,066
57742	Abhd1	abhydrolase domain containing 1	2,26	0,00068
68337	Crip2	cysteine rich protein 2	2,19	6,90E-05
223227	Sox21	SRY (sex determining region Y)-box 21	2,1	0,0098
20249	Scd1	stearoyl-Coenzyme A desaturase 1	2,03	0,001
15122		Mus musculus hemoglobin alpha, adult chain 1 (Hba-a1), mRNA.	1,92	0,19
19652	Rbm3	RNA binding motif protein 3	1,88	0,0014
14733	Gpc1	glypican 1	1,88	0,018
406217	Bex4	brain expressed X-linked 4	1,79	0,014
14472	Gbx2	gastrulation brain homeobox 2	1,78	0,019
223227	Sox21	SRY (sex determining region Y)-box 21	1,77	0,014
20350	Sema3f	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3F	1,67	0,014
192212	Prom2	prominin 2	1,65	0,048
69195	Tmem121	transmembrane protein 121	1,64	0,14
64293	Stk32b	serine/threonine kinase 32B	1,64	0,00092
23962	Oasl2	2'-5' oligoadenylate synthetase-like 2	1,57	0,0045
192236	Hps1	Hermansky-Pudlak syndrome 1	-2,05	0,00082
54150	Rdh7	retinol dehydrogenase 7	-2,08	0,033
233549	Mogat2	monoacylglycerol O-acyltransferase 2	-2,12	0,0023
69710	Arap1	ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 1	-2,12	0,0015
11522	Adh1	alcohol dehydrogenase 1 (class I)	-2,13	0,012
56018	Stard10	START domain containing 10	-2,14	0,0017
12780	Abcc2	ATP-binding cassette, sub-family C (CFTR/MRP), member 2	-2,18	0,0059
54150	Rdh7	retinol dehydrogenase 7	-2,23	0,032
56388	Cyp3a25	cytochrome P450, family 3, subfamily a, polypeptide 25	-2,4	0,034
11522	Adh1	alcohol dehydrogenase 1 (class I)	-2,44	0,0073
17921	Myo7a	myosin VIIA	-2,54	0,0032
13112	Cyp3a11	cytochrome P450, family 3, subfamily a, polypeptide 11	-2,82	0,065
17701	Msx1	msh homeobox 1	-3,05	0,064
233549	Mogat2	monoacylglycerol O-acyltransferase 2	-3,14	0,0037
233571	P2ry6	pyrimidinergic receptor P2Y, G-protein coupled, 6	-3,31	0,00024
18479	Pak1	p21 protein (Cdc42/Rac)-activated kinase 1	-3,4	0,003
68185	Coa4	cytochrome c oxidase assembly factor 4	-3,44	8,00E-04
52443	Mrpl48	mitochondrial ribosomal protein L48	-3,56	0,0018
52443	Mrpl48	mitochondrial ribosomal protein L48	-3,77	0,0016
27050	Rps3	ribosomal protein S3	-6,14	0,0025

increase in expression of the Wnt target genes *Ascl2*, *Axin2*, *Lgr5*, *Olfm4*, and *SP5* in Msx1-deficient cells compared to Msx1-proficient cells. On the other hand, expression of sucrose isomaltase (*SI*) and chromogranin A (*CHGA*), genes encoding markers of enterocytes and enteroendocrine cells, respectively, decreased (Figure 33). These observations further support our hypothesis that Msx1 loss leads to decreased differentiation of intestinal epithelial cells.

Subsequently, Msx1 expression and function was analyzed also in the colon. Analogically to the small intestinal epithelium, Msx1 mRNA and protein was not observed in the colon at physiological conditions. However, seven days after tamoxifen-induced Apc inactivation, Msx1 protein was detected in the upper parts of the hyperplastic crypts in the *Apc*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{+/+}$) colon. Although the recombination efficacy of the conditional *Apc* alleles along in the colon was similar to the small intestine (the conclusion was based on the magnitude of the colon crypt hyperplasia), the Msx1 staining was less pronounced. Moreover, Msx1 protein was stained predominantly in proximal third of the colon; in the distal portion of the tissue the *Msx1* gene was either not expressed or the expression level was below the limit of immunohistochemical detection. Simultaneous inactivation of *Apc* and *Msx1* in *Apc*$^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$) mice resulted in reduced numbers of

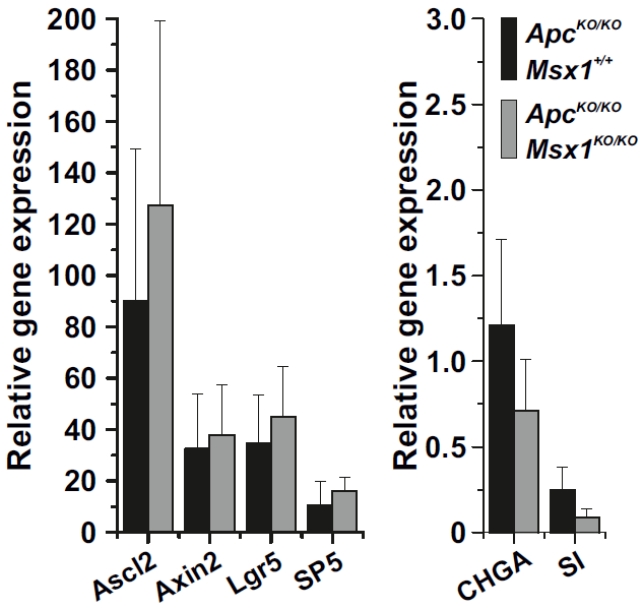


Figure 33 | Expression of cell differentiation markers decreases in Apc-/Msx1-double-deficient small intestine. Quantitative RT-PCR analysis in *Apc*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{+/+}$) and *Apc*$^{cKO/cKO}$ *Msx1*$^{cKO/cKO}$ *Villin-CreERT2* (*Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$) small intestine 7 days after tamoxifen administration; control samples were obtained from mice that were administered with the solvent only. Ct values were normalized to *β-actin* gene expression; *CHGA*, chromogranin A; *SI*, sucrose isomaltase. The diagram shows expression in *Apc*$^{KO/KO}$ *Msx1*$^{+/+}$ and *Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$ cells relative to wild-type cells (the gene expression level in control mice was arbitrarily set to 1). RNA samples obtained from three tamoxifen-treated mice of both strains and four control animals were analyzed; qRT-PCR reactions were run in technical triplicates; error bars indicate SDs. The fold change in *Apc*$^{KO/KO}$ *Msx1*$^{+/+}$ and *Apc*$^{KO/KO}$ *Msx1*$^{KO/KO}$ mice in comparison to control mice was in the *Msx1* mRNA levels 12337.82 and 1114.82, respectively, and in the *Msx2* mRNA levels 1523.82 and 1543.41. These results were not included in the diagrams due to the high values.

protiferating cells in the upper part of the crypts. Moreover, the absence of goblet cells marker mucin 2, indicated a loss of cell differentiation in both the Msx1-proficient and the Msxl-deficient epithelium (Figure 34).

In order to identify genes affected by Msx1 loss, expression profiling of epithelial cells isolated from proximal third of Apc-/Msx1-double deficient, Apc-deficient, and control colon was performed. Quite expectedly, expression of many genes differed significantly between control and Msx-deficient or Apc-/Msx1-double-deficient cells (Table 9). However, only the gene encoding serine/threonine kinase 32B (*Stk32b*) exhibited significantly different expression between Msx1-proficient and Msx1-deficient cells.

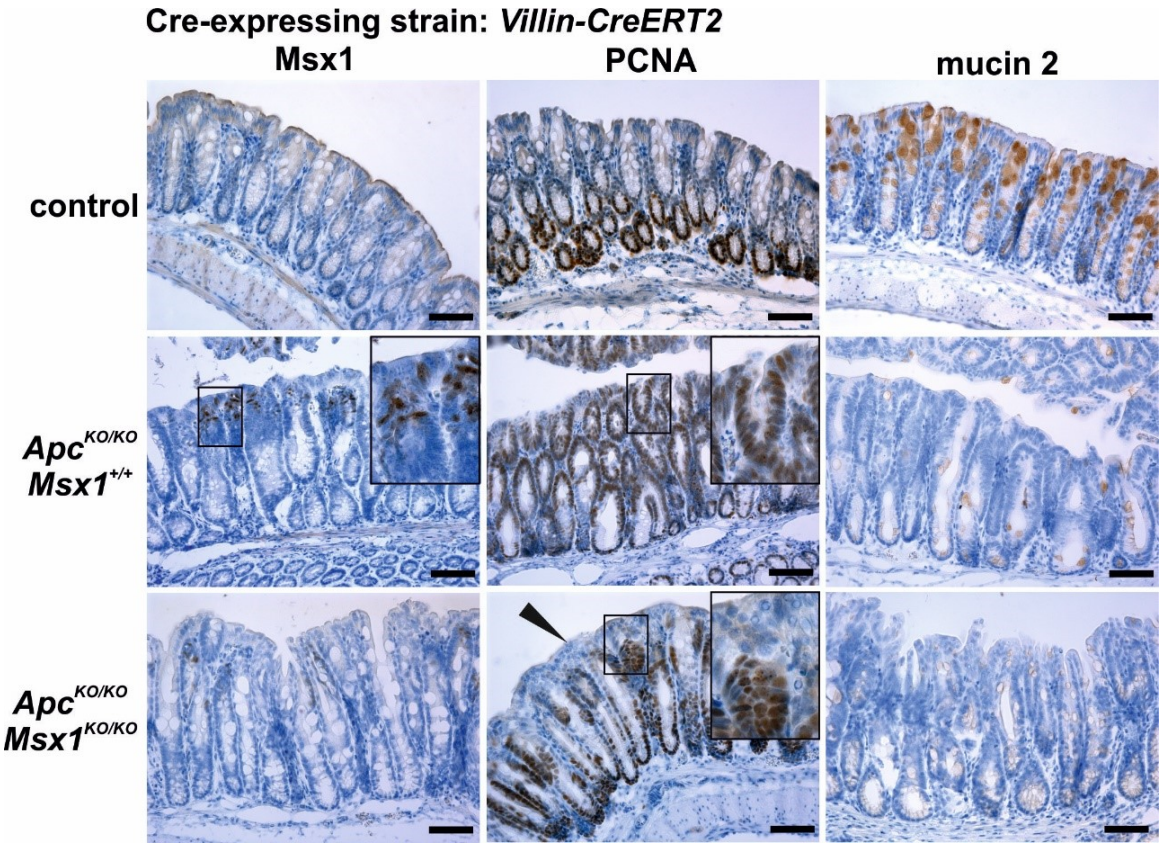


Figure 34 | Loss of Msx1 reduces cell differentiation in Apc-deficient colon.
Immunohistochemical staining of Msx1, PCNA, and mucin 2 (Muc2) in $Apc^{cKO/cKO}$ *Villin-CreERT2* ($Apc^{KO/KO}$ $Msx1^{+/+}$) and $Apc^{cKO/cKO}$ $Msx1^{cKO/cKO}$ *Villin-CreERT2* ($Apc^{KO/KO}$ $Msx1^{KO/KO}$) colon 7 days after tamoxifen administration; control animals were administered with the solvent only. Of note, PCNA staining was lost in upper crypt portions in the $Apc^{cKO/cKO}$ $Msx1^{cKO/cKO}$ colon epithelium. Nevertheless, the expression of goblet cells marker mucin 2 was reduced independently of the Msx1 status. Three animals of each genotype were analyzed, representative images are shown. The sections were counterstained with hematoxylin; boxed areas are magnified in the insets; scale bar: 0.15 mm.

Table 9 | Differentialy expressed genes (|logFC| ≥ 0.8) in the Apc-/Msx1-double-deficient colon mucosa when compared to Apc-deficient colon mucosa with intact Msx1.
The table shows differentially expressed genes in colon epithelium isolated from *Apc$^{cKO/cKO}$ Msx1$^{cKO/cKO}$ Villin-CreERT2* mice 7 days after tamoxifen administration (1 mg per animal); *Apc$^{cKO/cKO}$ Villin-CreERT2* mice were used as a control. The experiment was performed in four biological replicates.

PROBE ID	ENSEMBL ID	SYMBOL	GENENAME	logFC	p-value	q-value
ENSMUST00000094836	ENSMUSG00000029123	Stk32b	serine/threonine kinase 32B	2.06e+00	3.70e-09	0.000314
ENSMUST00000103399	ENSMUSG00000076598	Igkv3-7	immunoglobulin kappa variable 3-7	1.94e+00	8.93e-04	1.000000
ENSMUST00000197560	ENSMUSG00000076598	Igkv3-7	immunoglobulin kappa variable 3-7	1.94e+00	8.93e-04	1.000000
ENSMUST00000177591	ENSMUSG00000096768	Erdr1	erythroid differentiation regulator 1	1.48e+00	2.43e-02	1.000000
ENSMUST00000177671	ENSMUSG00000096768	Erdr1	erythroid differentiation regulator 1	1.44e+00	2.60e-02	1.000000
ENSMUST00000178789	ENSMUSG00000095562	Gm21887	predicted gene, 21887	1.37e+00	2.73e-02	1.000000
ENSMUST00000179483	ENSMUSG00000096768	Erdr1	erythroid differentiation regulator 1	1.37e+00	2.73e-02	1.000000
ENSMUST00000180251	ENSMUSG00000095562	Gm21887	predicted gene, 21887	1.37e+00	2.73e-02	1.000000
ENSMUST00000044159	ENSMUSG00000060807	Serpina6	serine (or cysteine) peptidase inhibitor, clade A, member 6	1.36e+00	1.54e-04	1.000000
ENSMUST00000179077	ENSMUSG00000096768	Erdr1	erythroid differentiation regulator 1	1.34e+00	2.64e-02	1.000000
ENSMUST00000100692	ENSMUSG00000095528	Gm10375	predicted gene 10375	1.16e+00	3.12e-02	1.000000
ENSMUST00000163970	ENSMUSG00000095528	Gm10375	predicted gene 10375	1.16e+00	3.12e-02	1.000000
ENSMUST00000196706	ENSMUSG00000027869	Hsd3b6	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 6	1.15e+00	4.68e-04	1.000000
ENSMUST00000211636	ENSMUSG00000040640	Erc2	ELKS/RAB6-interacting/CAST family member 2	1.15e+00	8.61e-04	1.000000
ENSMUST00000144418	ENSMUSG00000028469	Npr2	natriuretic peptide receptor 2	1.15e+00	1.48e-02	1.000000
ENSMUST00000172766	ENSMUSG00000050423	Ppp1r3g	protein phosphatase 1, regulatory (inhibitor) subunit 3G	1.13e+00	5.97e-03	1.000000
ENSMUST00000113512	ENSMUSG00000073643	Wdfy1	WD repeat and FYVE domain containing 1	1.12e+00	3.83e-03	1.000000
ENSMUST00000113513	ENSMUSG00000073643	Wdfy1	WD repeat and FYVE domain containing 1	1.12e+00	3.83e-03	1.000000
ENSMUST00000113514	ENSMUSG00000073643	Wdfy1	WD repeat and FYVE domain containing 1	1.12e+00	3.83e-03	1.000000
ENSMUST00000113515	ENSMUSG00000073643	Wdfy1	WD repeat and FYVE domain containing 1	1.12e+00	3.83e-03	1.000000
ENSMUST00000187005	ENSMUSG00000073643	Wdfy1	WD repeat and FYVE domain containing 1	1.08e+00	3.73e-03	1.000000
ENSMUST00000203150	ENSMUSG00000030361	Klrb1a	killer cell lectin-like receptor subfamily B member 1A	1.06e+00	4.21e-02	1.000000
ENSMUST00000172486	ENSMUSG00000015222	Map2	microtubule-associated protein 2	1.05e+00	8.34e-03	1.000000
ENSMUST00000135885	ENSMUSG00000029095	Ablim2	actin-binding LIM protein 2	1.04e+00	1.43e-04	1.000000
ENSMUST00000186702	ENSMUSG00000041460	Cacna2d4	calcium channel, voltage-dependent, alpha 2/delta subunit 4	1.03e+00	5.87e-03	1.000000
ENSMUST00000131920	ENSMUSG00000023267	Gabrr2	gamma-aminobutyric acid (GABA) C receptor, subunit rho 2	1.03e+00	3.48e-02	1.000000
ENSMUST00000171262	ENSMUSG00000006711	D130043K22Rik	RIKEN cDNA D130043K22 gene	1.01e+00	1.44e-02	1.000000
ENSMUST00000186394	ENSMUSG00000074109	Mrgprx2	MAS-related GPR, member X2	1.00e+00	8.34e-03	1.000000
ENSMUST00000103483	ENSMUSG00000076674	Ighv3-8	immunoglobulin heavy variable V3-8	9.86e-01	3.05e-01	1.000000
ENSMUST00000185329	ENSMUSG00000025932	Eya1	EYA transcriptional coactivator and phosphatase 1	9.84e-01	4.82e-02	1.000000
ENSMUST00000040361	ENSMUSG00000039347	Atp6v0e2	ATPase, H+ transporting, lysosomal V0 subunit E2	9.79e-01	5.14e-03	1.000000
ENSMUST00000136987	ENSMUSG00000043587	Pxylp1	2-phosphoxylose phosphatase 1	9.68e-01	9.28e-04	1.000000
ENSMUST00000144697	ENSMUSG00000026999	Nup35	nucleoporin 35	9.67e-01	1.95e-02	1.000000
ENSMUST00000153129	ENSMUSG00000028047	Thbs3	thrombospondin 3	9.65e-01	4.39e-02	1.000000
ENSMUST00000103350	ENSMUSG00000076549	Igkv4-68	immunoglobulin kappa variable 4-68	9.63e-01	1.02e-02	1.000000
ENSMUST00000137290	ENSMUSG00000031698	Mylk3	myosin light chain kinase 3	9.62e-01	1.37e-02	1.000000
ENSMUST00000169797	ENSMUSG00000037849	Ifi206	interferon activated gene 206	9.49e-01	7.66e-02	1.000000
ENSMUST00000155275	ENSMUSG00000021596	Mctp1	multiple C2 domains, transmembrane 1	9.39e-01	6.10e-03	1.000000
ENSMUST00000162154	ENSMUSG00000022148	Fyb	FYN binding protein	9.37e-01	1.19e-02	1.000000
ENSMUST00000161947	ENSMUSG00000022148	Fyb	FYN binding protein	9.29e-01	7.54e-03	1.000000
ENSMUST00000190151	ENSMUSG00000021209	Ppp4r4	protein phosphatase 4, regulatory subunit 4	9.08e-01	7.42e-03	1.000000
ENSMUST00000172478	ENSMUSG00000074369	Obox2	oocyte specific homeobox 2	8.92e-01	1.01e-02	1.000000
ENSMUST00000174076	ENSMUSG00000074369	Obox2	oocyte specific homeobox 2	8.92e-01	1.01e-02	1.000000
ENSMUST00000174305	ENSMUSG00000074369	Obox2	oocyte specific homeobox 2	8.92e-01	1.01e-02	1.000000
ENSMUST00000149336	ENSMUSG00000029651	Mtus2	microtubule associated tumor suppressor candidate 2	8.84e-01	5.44e-03	1.000000
ENSMUST00000194041	ENSMUSG00000026587	Astn1	astrotactin 1	8.62e-01	1.16e-04	1.000000
ENSMUST00000213557	ENSMUSG00000071317	Bves	blood vessel epicardial substance	8.61e-01	1.04e-02	1.000000
ENSMUST00000015576	ENSMUSG00000022226	Mcpt2	mast cell protease 2	8.59e-01	7.10e-04	1.000000
ENSMUST00000207685	ENSMUSG00000035177	Nlrp2	NLR family, pyrin domain containing 2	8.50e-01	6.82e-04	1.000000
ENSMUST00000204277	ENSMUSG00000039347	Atp6v0e2	ATPase, H+ transporting, lysosomal V0 subunit E2	8.50e-01	3.72e-03	1.000000
ENSMUST00000201736	ENSMUSG00000094719	Gm5108	predicted gene 5108	8.48e-01	1.46e-03	1.000000
ENSMUST00000135355	ENSMUSG00000021645	Smn1	survival motor neuron 1	8.46e-01	4.36e-02	1.000000
ENSMUST00000142251	ENSMUSG00000051747	Ttn	titin	8.43e-01	8.03e-02	1.000000
ENSMUST00000195849	ENSMUSG00000034837	Gnat1	guanine nucleotide binding protein, alpha transducing 1	8.41e-01	2.01e-02	1.000000
ENSMUST00000202984	ENSMUSG00000006641	Slc5a6	solute carrier family 5 (sodium-dependent vitamin transporter), member 6	8.29e-01	1.50e-02	1.000000
ENSMUST00000022836	ENSMUSG00000022227	Mcpt1	mast cell protease 1	8.28e-01	1.36e-02	1.000000
ENSMUST00000176196	ENSMUSG00000032595	Cdhr4	cadherin-related family member 4	8.25e-01	1.70e-02	1.000000
ENSMUST00000177093	ENSMUSG00000032595	Cdhr4	cadherin-related family member 4	8.25e-01	1.70e-02	1.000000
ENSMUST00000141085	ENSMUSG00000041216	Clvs1	clavesin 1	8.23e-01	3.49e-02	1.000000
ENSMUST00000095450	ENSMUSG00000071178	Serpina1b	serine (or cysteine) preptidase inhibitor, clade A, member 1B	8.20e-01	2.66e-02	1.000000
ENSMUST00000164454	ENSMUSG00000071178	Serpina1b	serine (or cysteine) preptidase inhibitor, clade A, member 1B	8.20e-01	2.66e-02	1.000000
ENSMUST00000186166	ENSMUSG00000071178	Serpina1b	serine (or cysteine) preptidase inhibitor, clade A, member 1B	8.20e-01	2.66e-02	1.000000
ENSMUST00000195095	ENSMUSG00000104098	AA619741	expressed sequence AA619741	8.18e-01	3.66e-02	1.000000
ENSMUST00000168044	ENSMUSG00000043557	Mdga1	MAM domain containing glycosylphosphatidylinositol anchor 1	8.15e-01	2.99e-04	1.000000
ENSMUST00000189541	ENSMUSG00000016918	Sulf1	sulfatase 1	8.13e-01	1.79e-02	1.000000
ENSMUST00000172308	ENSMUSG00000072731	Gm3715	predicted gene 3715	8.08e-01	3.60e-03	1.000000
ENSMUST00000190082	ENSMUSG00000026246	Alppl2	alkaline phosphatase, placental-like 2	8.08e-01	1.82e-02	1.000000
ENSMUST00000148715	ENSMUSG00000009246	Trpm5	transient receptor potential cation channel, subfamily M, member 5	8.06e-01	4.03e-02	1.000000
ENSMUST00000191403	ENSMUSG00000099826	Scgb2b10	secretoglobin, family 2B, member 10	8.04e-01	1.00e-01	1.000000
ENSMUST00000103323	ENSMUSG00000076522	Igkv16-104	immunoglobulin kappa variable 16-104	8.00e-01	1.14e-01	1.000000

Quantitative RT-PCR analysis also showed that only *Msx1* and *Stk32B* expression levels significantly changed due to the *Msx1* gene inactivation (Figure 35). The gene encoding *SRY-related HMG box transcription factor 17* (*Sox17*) was previously described as the Wnt target gene[63] and a regulator of the canonical Wnt signaling[335, 428]. The qRT-PCR analysis revealed elevated *Sox17* mRNA levels in Apc-/Msx1-double-deficient epithelial cells in comparison to Apc-deficient cells (Figure 36A). *Sox17* mRNA expression was evaluated by *in situ* hybridization at sections obtained from $Apc^{cKO/cKO}$ $Msx1^{cKO/cKO}$ *Villin-CreERT2* and $Apc^{cKO/cKO}$ $Msx1^{cKO/cKO}$ *Lgr5-EGFP-IRES-CreERT2* mice sacrificed 7 and 14 days upon tamoxifen administration, respectively. In $Apc^{cKO/cKO}$ $Msx1^{cKO/cKO}$ *Villin-CreERT2* mice, *Sox17* mRNA was distinctly present in the villi (red arrowheads) and displayed a much stronger signal than in $Apc^{cKO/cKO}$ *Villin-CreERT2*. In $Apc^{cKO/cKO}$

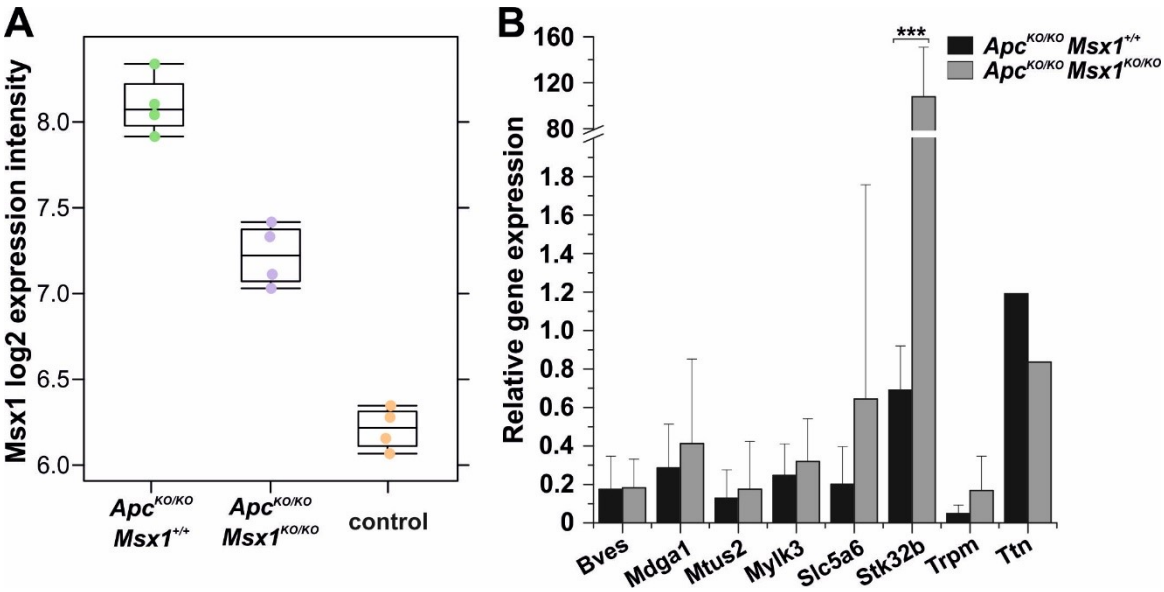


Figure 35 | Gene expression profiling analysis of the Apc-deficient and Apc-/Msx1-double-deficient colon epithelium.
Total RNA was isolated from epithelial cells isolated from proximal colon of $Apc^{cKO/cKO}$ *Villin-CreERT2* ($Apc^{KO/KO}$ $Msx1^{+/+}$) and $Apc^{cKO/cKO}$ $Msx1^{cKO/cKO}$ *Villin-CreERT2* ($Apc^{KO/KO}$ $Msx1^{KO/KO}$) mice 7 days after tamoxifen administration; control samples were obtained from animals that were administered with the solvent only. (A) Results of the DNA microarray analysis for *Msx1* gene expression. The boxed areas correspond to the second and third quartiles; the range of the values is given by "whiskers" above and below each box; the median is indicated by the crossline. (B) Quantitative RT-PCR analysis of selected genes from the expression profiling analysis confirmes significantly increased expression of the *Stk32B* gene. The diagram shows expression in $Apc^{KO/KO}$ $Msx1^{+/+}$ and $Apc^{KO/KO}$ $Msx1^{KO/KO}$ cells relative to wild-type cells (the gene expression level in control mice was arbitrarily set to 1). The fold change in $Apc^{KO/KO}$ $Msx1^{+/+}$ and $Apc^{KO/KO}$ $Msx1^{KO/KO}$ mice in comparison to control mice was in the *Msx1* mRNA levels 862.66 and 76.74, respectively, and in the *Msx2* mRNA levels 840.45 and 1344.64. These results were not included in the diagrams due to the high values. The qRT-PCR reactions were run in technical triplicates; error bars indicate SDs; ***, $p < 0.001$.

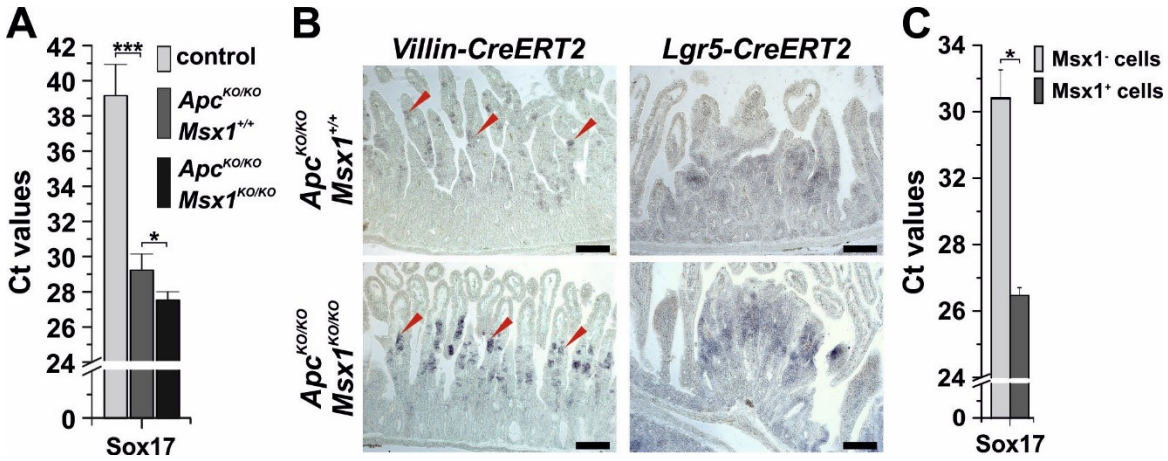


Figure 36 | Analysis of *Sox17* expression in mouse intestinal epithelial cells.
(A) Quantitative RT-PCR analysis of *Sox17* expression in the $Apc^{cKO/cKO}$ *Villin-CreERT2* ($Apc^{KO/KO}$ $Msx1^{+/+}$) and $Apc^{cKO/cKO}$ $Msx1^{cKO/cKO}$ *Villin-CreERT2* ($Apc^{KO/KO}$ $Msx1^{KO/KO}$) small intestine 7 days after tamoxifen administration; control samples were obtained from mice that were administered with the solvent only. Ct values were normalized to *β-actin* gene expression; RNA samples were obtained from three animals of all genotypes; qRT-PCR reactions were run in technical triplicates; error bars indicate SDs; *, $p < 0.05$; ***, $p < 0.001$. (B) Detection of *Sox17* transcript in crypt hyperplasia and early adenomas developed in the $Apc^{KO/KO}$ $Msx1^{+/+}$ and $Apc^{KO/KO}$ $Msx1^{KO/KO}$ epithelium using *in situ* hybridization; tissues were isolated 7 or 14 days after tamoxifen administration, respectively. In the *Villin-CreERT2* strain, anti-*Sox17* probe stained pouches of cells on villi (red arrowheads). In the *Lgr5-CreERT2* strain, the signal was detected in adenomas. Scale bar: 0.3 mm. (C) Quantitative RT-PCR analysis revealed significantly increased expression of *Sox17* in Msx1-positive (ectopic crypt) cells isolated from $Apc^{cKO/cKO}$ *Villin-CreERT2* small intestinal epithelium 7 days upon Apc depletion. The total RNA level was normalized to expression of internal housekeeping gene *β-actin* that was arbitrarily set to 23. Error bars indicate SDs; *, $p < 0.05$.

$Msx1^{cKO/cKO}$ *Lgr5-EGFP-IRES-CreERT2* intestines the *Sox17* mRNA was clearly detected in (villous) adenomas and weakly in the (tubular) adenomas of $Apc^{cKO/cKO}$ *Lgr5-EGFP-IRES-CreERT2* (Figure 36B). Finally, *Sox17* mRNA levels were significantly increased in Msx1-positive epithelial cells obtained from the $Apc^{cKO/cKO}$ *Villin-CreERT2* small intestinal epithelium 7 days upon tamoxifen administration (i. e. in ectopic crypt cells; Figure 36C). The previous study suggested that Sox17 promotes degradation of β-catenin and Tcf4[335]. Since the *Sox17* increase was observed in Apc-/Msx1-double deficient intestines, we also tested whether the amount of β-catenin and Tcf4 proteins were altered. However, no significant changes in the β-catenin and Tcf4 protein levels or distribution were observed (Figure 37).

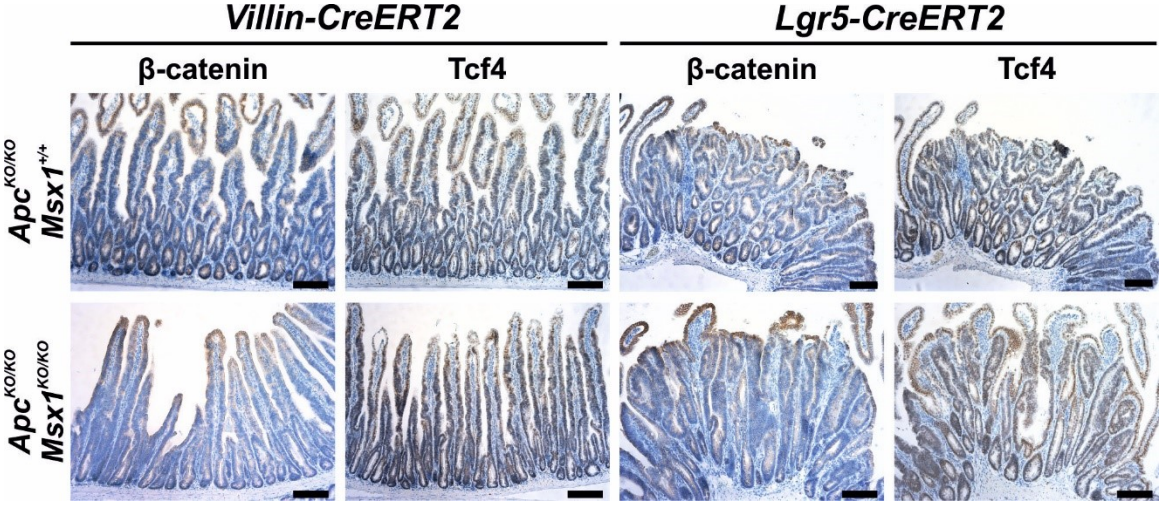


Figure 37 | Analysis of β-catenin and Tcf4 expression in Msx1-deficient small intestine. Immunohistochemical stainig of β-catenin and Tcf4 proteins in crypt hyperplasia and early adenomas developed in $Apc^{KO/KO}\,Msx1^{+/+}$ and $Apc^{KO/KO}\,Msx1^{KO/KO}$ epithelium 7 or 14 days upon tamoxifen administration (1 mg per animal), respectively; Villin- and Lgr5- drivers were used for CreERT2 expression. No difference in the staining was observed between Msx1-deficient and Msx1-proficient tissue. Three animals of each genotype were analyzed; representative images are shown. The sections were counterstained with hematoxylin; scale bar: 0.3 mm.

4.3 The MSX1 function in human colorectal cancer

MSX1 expression in human tissues was ascertained using the BioGPS portal. Gene expression analysis revealed that in human tissues, *MSX1* is predominantly espressed in samples obtained from colorectal adenocarcinoma (Figure 38). To verify this finding

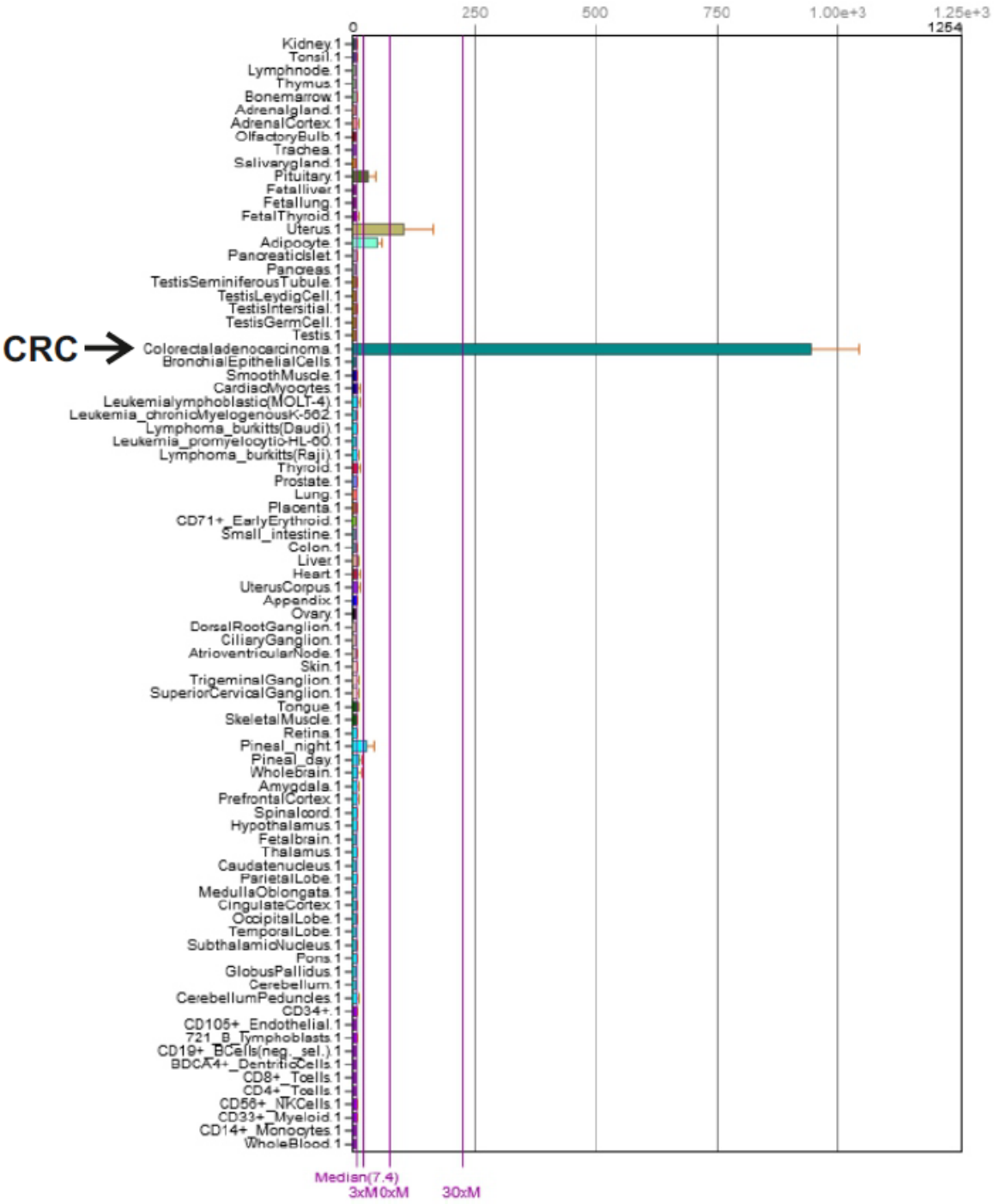


Figure 38 | MSX1 is abundant in human colorectal adenocarcinoma.
Gene expression profiling of 79 human tissues provided by the BioGPS portal (dataset: GeneAtlas U133A, gcrma; probeset: 205932_s_at). The diagram shows *MSX1* gene expression in human tissues with a remarkable abundance in colorectal adenocarcinomas.

obtained from publically available dataset, we collected RNA samples isolated from human colorectal neoplasia specimens and matching healthy mucosa. Quantitative RT-PCR analysis revealed increased *MSX1* mRNA levels in all stages of intestinal neoplasia tested. However, *MSX1* displayed the highest upregulation in adenomas with low grade dysplasia and the mRNA levels exhibited decreasing tendency with progression to more advanced neoplastic stages (Figure 39A). Histological analysis of mouse tumors arising from Msx1-deficient cells display altered morphology at later stages of development reminding more progressed villous adenomas. In humans, such morphological conversion of colorectal neoplasia indicates tumor progression associated with elevated risk of malignancy. Therefore *MSX1* expression was further analyzed in colorectal tumors with tubular, tubulovillous or villous morphology. However, no correlation between tumor morphology and the *MSX1* expression levels was observed (Figure 39B).

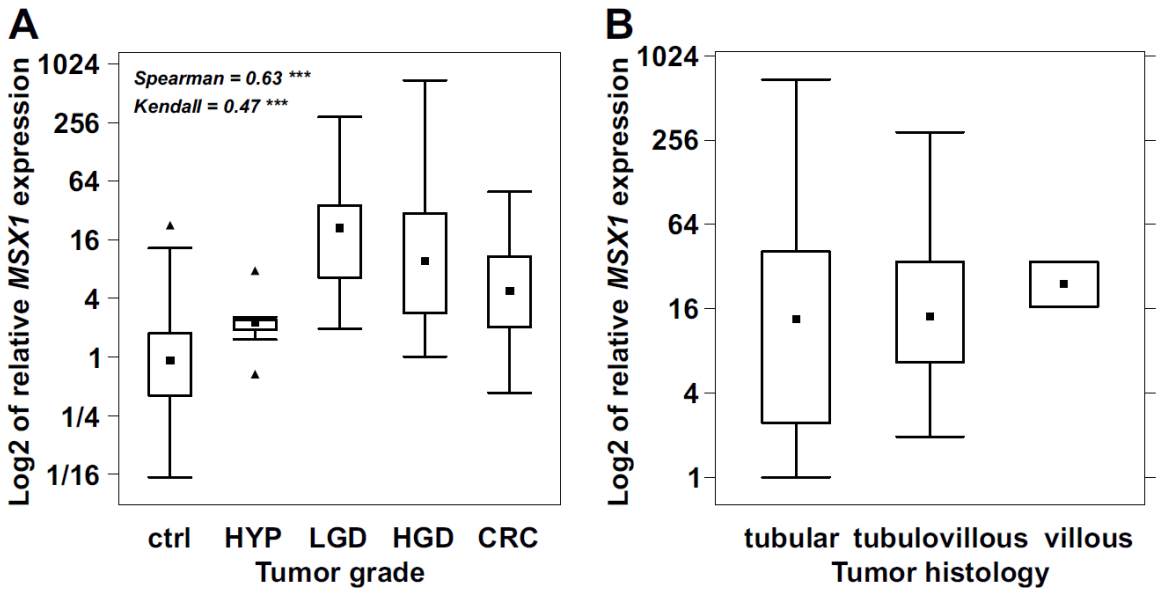


Figure 39 | *MSX*1 expression is abundant in human colorectal tumors.
Quantitative RT-PCR analysis of *MSX1* mRNA in samples obtained from human colorectal neoplasia specimens. The samples were divided into categories based on the tumor grade (A) or tumor histology (B) as follows: (A) healthy tissue (ctrl), hyperplastic adenomas (HYP; n = 9), adenomas displaying low-grade dysplasia (LGD; n = 27) or high-grade dysplasia (HGD; n = 24), and colorectal carcinoma (CRC; n = 12); (B) tubular (n = 20), tubulovillous (n = 31), and villous (n = 3). The amounts of RNA in individual isolates were normalized to the geometric average of Ct values of housekeeping genes *UBB* and *β2-microglobulin*. The boxed areas correspond to the second and third quartiles; the median of ΔCt values for each group of samples is indicated as the black square. The range of the values is given by "whiskers" above and below each box; outliers are indicated by black triangles. The significance between the *MSX1* mRNA level and neoplasia progression is demonstrated by the Spearman ($\rho = 0.63$) and Kendall ($\tau = 0.47$) coefficient values; ***, $p < 0.001$.

Generation and analysis of MSX1-deficient CRC cells

In order to select suitable cells for following experiments, MSX1 expression was tested in several human cell lines. The highest level of *MSX1* mRNA was detected in SW480 and SW620 cells; moreover, MSX1 in these cells also displayed a distinct protein staining using western blot (Figure 40). Note that SW480 cells were derived from a primary tumor and SW620 from a lymph node metastasis of the same patient[186].

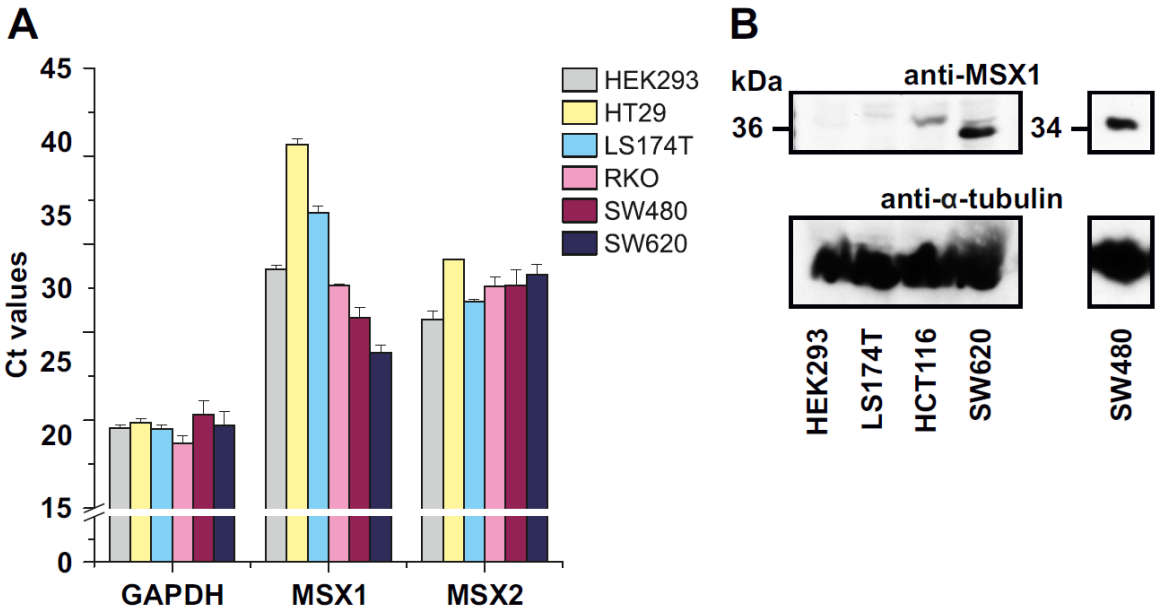


Figure 40 | MSX1 expression in human cells. (A) Quantitative RT-PCR analysis of *MSX1* expression in RNA samples obtained from human embryonic kidney cells (HEK293) and human CRC cell lines HT29, LS174T, RKO, SW480, and SW620. The diagram shows total RNA levels normalized to expression of internal housekeeping gene *β-actin* that was arbitrarily set to 17. The qRT-PCR reactions were run in technical triplicates; error bars indicate SDs. (B) Western blotting of lysates obtained from HEK293 and indicated human CRC cells with anti-MSX1 antibody confirmed production of MSX1 protein in SW480 and SW620 cells. Anti-α-tubulin antibody was used as a loading control.

To examine MSX1 function in human CRC cells, *MSX1* gene was disrupted in SW620 cells using the CRISPR/Cas9-mediated gene targeting. Two sites in the exon 1 were targeted with different guide RNAs (gRNAs), control cells were transduced with a lentiCRISPR vector encoding no gRNA (Figure 41A). Quantitative RT-PCR analysis confirmed *MSX1* loss and a negligible upregulation of *MSX2* in MSX1-deficient cells. As it was previously reported that MSX1 can regulate the Wnt signaling levels[283, 363], several Wnt target genes were included in the analysis. Whereas *TNFRSF19* (alias *TROY*) expression remained unchanged, other Wnt target genes *ASCL2* and *AXIN2* exhibited

elevated mRNA expression levels in MSX1-deficient cells (Figure 41B). Western blot analysis confirmed decrease in MSX1 protein levels in individual clones (Figure 41C).

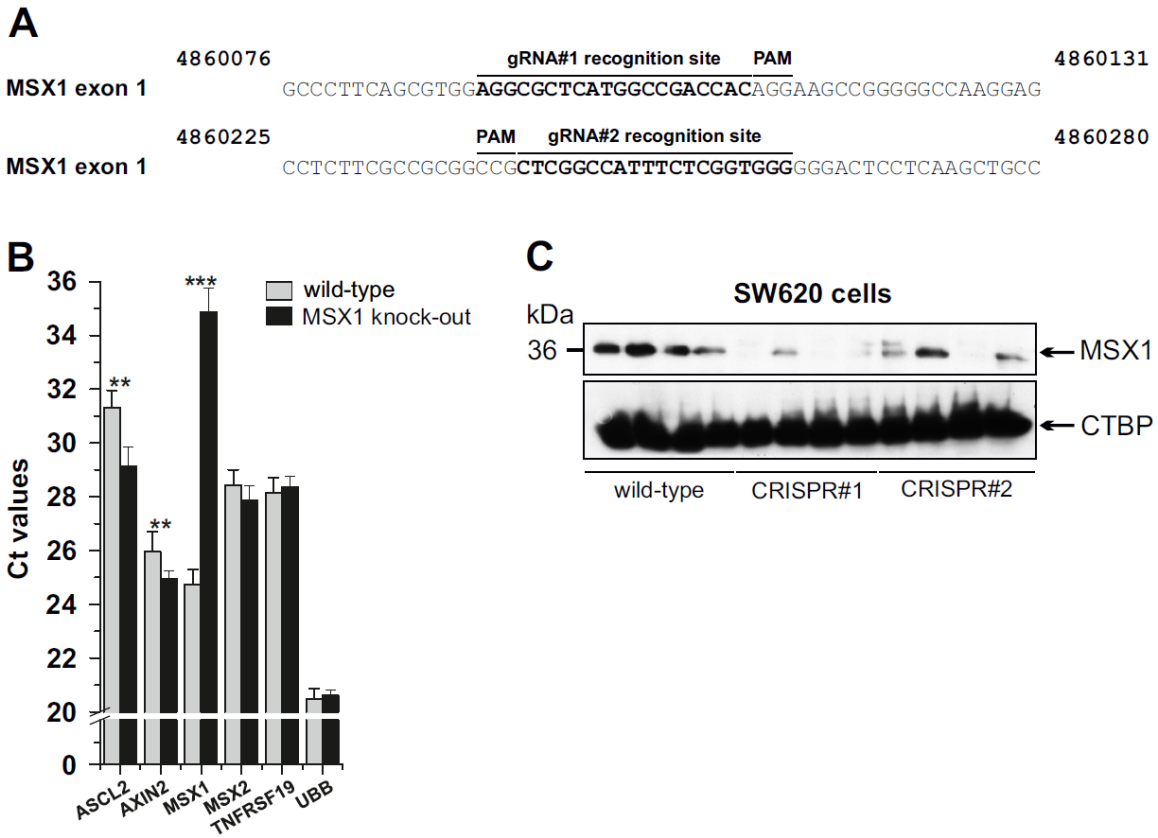


Figure 41 | Generation and analysis of SW620 cells with the disrupted *MSX1* gene. (A) DNA sequences display guide RNA (gRNA) recognition sites (in bold) in the first exon of the *MSX1* gene; PAM sequence is overlined. The numbers above sequences indicate nucleotide positions in the human genome assembly GRCh38:CM000666.2. (B) Quantitative RT-PCR analysis of RNA obtained from SW620 single cell clones harboring truncated *MSX1* gene ("knock-out"; n = 9) confirms significant downregulation of *MSX1* mRNA; control samples were obtained from cells with intact *MSX1* ("wild-type" n = 3). The diagram shows average values of total RNA levels in *MSX1* wild-type and knock-out cell clones normalized to expression of internal housekeeping gene *β-actin* (arbitrarily set to 17). The qRT-PCR reactions were run in technical triplicates; error bars indicate SDs; **, $p < 0.01$, ***, $p < 0.001$. (C) Western blotting of lysates obtained from *MSX1* wild-type and knock-out cell clones with an anti-MSX1 antibody indicates decreas of the MSX1 protein amount. Four different cell clones are shown for both gRNAs (i.e. CRISPR#1 and CRISPR#2). An anti-α-tubulin antibody was used as a loading control.

The MSX1-deficient clones were viable and appeared to grow in culture slightly faster than the control clones. To test whether *MSX1* downregulation affected the cell cycle progression, SW620 wild-type (control) cells and cells harboring *MSX1* truncations (CRISPR#1 and CRISPR#2) were fixed and stained with propidium iodide (PI) solution and the cell cycle was analyzed by flow cytometer (Figure 42A). In addition, a viability

assay was performed to measure the proliferation of individual cell clones (Figure 42B). However, in both analyses MSX1-deficient clones did not exhibit change in cell cycle progression or proliferation rate. As previously published studies reported that MSX1 can regulate epithelial-mesenchymal transition[128, 299, 303], which is a process closely associated with changes in cellular movement, we next analyzed the ability of individual clones to repair a scratch in a confluent cell layer (i.e. the "wound healing assay"). Individual clones displayed slight differences in size of the "healed area", however, there was no significant difference between MSX1-deficient and control cells (Figure 42C). Finally, the clones were subcutaneously injected into the lumbar back area of immunodeficient NOD/SCID/GAMMA (NSG) mice to test their capabilities to grow as xenografts. The mice were sacrificed four weeks after injection and tumors were weighed, embedded in paraffin, and sections were analyzed by immunohistochemical staining. However, no differences in the tumor size were observed (Figure 42D, E). Moreover, MSX1-deficient tumors did not exhibit altered staining for a Ki-67 proliferating cells marker or p21 cell cycle inhibitor (Figure 43).

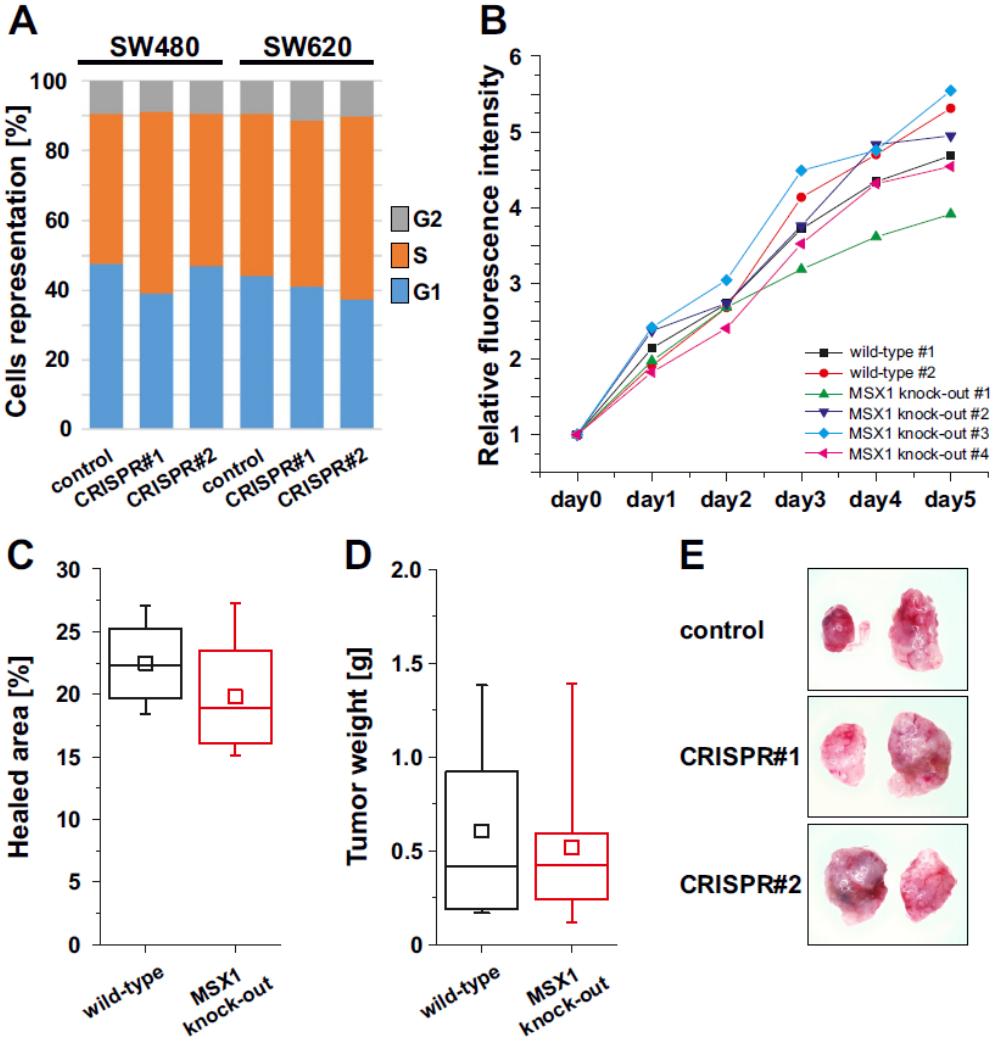


Figure 42 | MSX1-deficient SW620 cell clones do not display an altered proliferation rate or capacity to form xenograft. (A) SW480 and SW620 cells harboring truncations in the *MSX1* gene were isolated, fixed, and stained with propodium iodide (PI) solution. The diagram shows percentage of cells in cell cycle phases measured by flow cytometer. The experiment was performed twice, representative results are shown. (B) Cell viability assay performed on MSX1-deficient cell clones does not reveal any differences in the proliferation rate compared to wild-type cells. The diagram shows relative fluorescence intensity of the alamarBlue metabolic product. The measurements were performed in technical triplicates and average values of fluorescent intensity at indicated timepoints were normalized to day 0. The experiment was performed twice, representative results are shown. (C) The wound healing assay does not reveal any differences between MSX1-deficient (n = 3) and wild-type (n = 2) cell clones in terms of migration. The diagram shows percentage of the scraped area that was regrown ("healed") 22 hours after the scrape was created. The experiment was performed in six technical replicates (the boxes includes 12 and 18 values, respectively). The boxed areas correspond to the second and third quartiles; the median and mean value is indicated by a crossline or an empty square, respectively. The range of the values is given by "whiskers" above and below each box. (D) MSX1-deficient (n = 3) and wild-type (n = 3) cell clones were subcutaneously injected in the lumbar back area of NOD/SCID/GAMMA (NSG) mice (1×10^7 cells in 100 µl PBS per animal), 3 animals were used for each cell clone. Mice were sacrificed after 28 days, the tumors were resected and weighed. The experiment was performed twice, the diagram shows tumor weights combined from both experiments. The boxed areas correspond to the second and third quartiles; the median and mean value is indicated by a crossline or an empty square, respectively. The range of the values is given by "whiskers" above and below each box. (E) Macroscopic pictures of resected tumors derived from wild-tyxpe of MSX1 knock-out (CRISPR#1 and CRISPR#2) SW620 cell clones; representative images are shown.

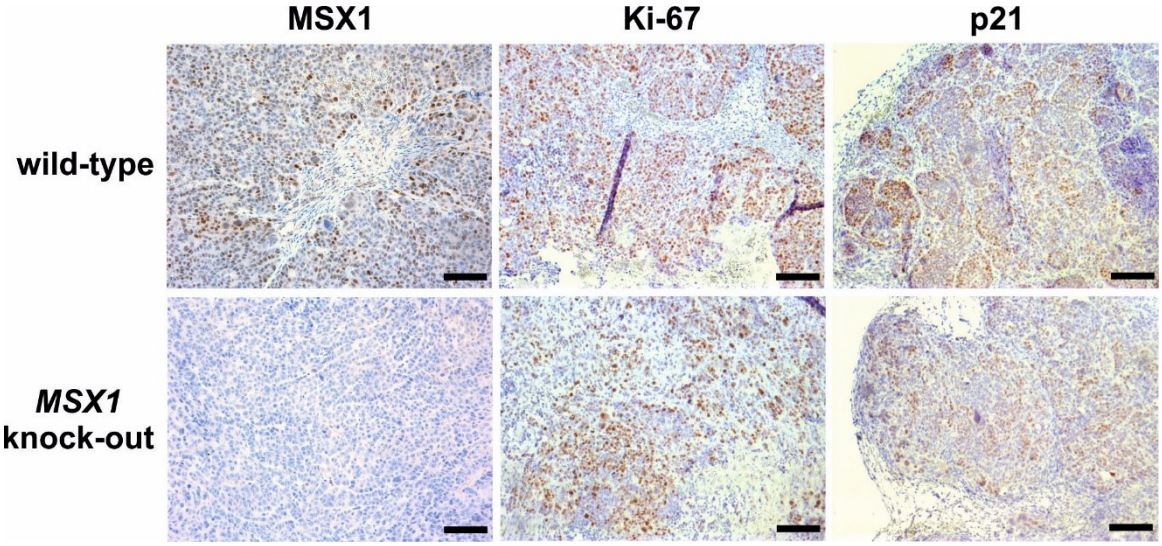


Figure 43 | MSX1-deficient tumors grown as xenografts do not display altered histology. Tumors derived from wild-type SW620 single cell clones (n = 3) and clones harboring truncations in the *MSX1* gene (n = 3) were embedded in paraffin and sectiones were immunohistochemically stained. Ki-67 and p21 proteins were detected using corresponding antibodies (brown signal); MSX1 protein was detected only in tumors derived from wild-type SW620 cells. Representative images are shown. The sections were counterstained with hematoxylin; scale bar: 0.3 mm.

Expression profiling of MSX1-deficient CRC cells

In order to identify MSX1-regulated genes in CRC cells, we performed expression profiling of MSX1-deficient SW620 cells. The profiling revealed 202 genes (including *ASCL2* and *SP5)* with significantly altered expression compared to control cells with intact *MSX1*; list of twenty genes with the most increased or decreased expression and fulfilling the signicifance criteria $|\log FC| \geq 1$ and q-value < 0.05 in MSX1-deficient SW620 cells is given in Table 10 (for a complete list of differentially expressed genes, see reference[114]).

Table 10 | Differentialy expressed genes ($|\log FC| \geq 1$; q-value < 0.05) in SW620 cells with the disrupted *MSX1* gene compared to SW620 cells with intact *MSX1*.

ENTREZ	SYMBOL	GENENAME	logFC	p-value
25984	KRT23	keratin 23, type I	4,99	3.4e-08
3860	KRT13	keratin 13, type I	4,12	3.6e-05
11009	IL24	interleukin 24	3,65	9.7e-07
430	ASCL2	achaete-scute family bHLH transcription factor 2	3,51	7.6e-07
2706	GJB2	gap junction protein, beta 2, 26kDa	3,27	4.4e-07
79083	MLPH	melanophilin	3,11	3.2e-05
9289	ADGRG1	adhesion G protein-coupled receptor G1	3,05	2.3e-07
54923	LIME1	Lck interacting transmembrane adaptor 1	3,01	0,00
54843	SYTL2	synaptotagmin-like 2	2,99	0,00
1473	CST5	cystatin D	2,93	5.1e-05
4923	NTSR1	neurotensin receptor 1 (high affinity)	2,87	0,00
56937	PMEPA1	prostate transmembrane protein, androgen induced 1	2,80	3.8e-05
80206	FHOD3	formin homology 2 domain containing 3	2,77	6.8e-06
9289	ADGRG1	adhesion G protein-coupled receptor G1	2,71	3.7e-06
4843	NOS2	nitric oxide synthase 2, inducible	2,65	2.2e-08
8771		Homo sapiens tumor necrosis factor receptor superfamily, member 6b, decoy	2,64	3.5e-05
8771		Homo sapiens tumor necrosis factor receptor superfamily, member 6b, decoy	2,63	1.3e-05
56937	PMEPA1	prostate transmembrane protein, androgen induced 1	2,60	2.9e-05
4071	TM4SF1	transmembrane 4 L six family member 1	2,44	7.1e-05
124056	NOXO1	NADPH oxidase organizer 1	2,44	3.9e-05
55244	SLC47A1	solute carrier family 47 (multidrug and toxin extrusion), member 1	-1,66	9.2e-05
5159	PDGFRB	platelet-derived growth factor receptor, beta polypeptide	-1,66	2.6e-05
343990	KIAA1211L	KIAA1211-like	-1,66	2.7e-06
30846	EHD2	EH-domain containing 2	-1,66	3.4e-05
343990	KIAA1211L	KIAA1211-like	-1,68	2.7e-07
2192	FBLN1	fibulin 1	-1,74	0,00
162494	RHBDL3	rhomboid, veinlet-like 3 (Drosophila)	-1,76	5.2e-05
146850	PIK3R6	phosphoinositide-3-kinase, regulatory subunit 6	-1,85	0,00
641700	ECSCR	endothelial cell surface expressed chemotaxis and apoptosis regulator	-1,86	0,00
3689	ITGB2	integrin, beta 2 (complement component 3 receptor 3 and 4 subunit)	-1,89	9.6e-05
6448	SGSH	N-sulfoglucosamine sulfohydrolase	-1,93	4.7e-06
5654	HTRA1	HtrA serine peptidase 1	-1,93	0,00
5138	PDE2A	phosphodiesterase 2A, cGMP-stimulated	-2,09	0,00
3851	KRT4	keratin 4, type II	-2,15	0,00
946	SIGLEC6	sialic acid binding Ig-like lectin 6	-2,37	1.3e-06
6280	S100A9	S100 calcium binding protein A9	-2,40	0,00
2018	EMX2	empty spiracles homeobox 2	-2,58	0,00
2018	EMX2	empty spiracles homeobox 2	-2,76	0,00
1277	COL1A1	collagen, type I, alpha 1	-2,87	0,00
24141	LAMP5	lysosomal-associated membrane protein family, member 5	-2,90	2.1e-07

Subsequent analysis using the online tool Enrichr[123, 181] did not reveal any biological process, signaling pathway, or molecular mechanism typical for the MSX1-deficient cells. However, we identified an overlap between our set of 202 genes and a set of 162 β-catenin-activated genes in SW480 cells identified by anti-β-CATENIN chromatin immunoprecipitation (ChIP)-sequencing (ChIP-seq)[400]. Interestingly, all overlapping genes were upregulated in MSX1-deficient cells (Figure 44).

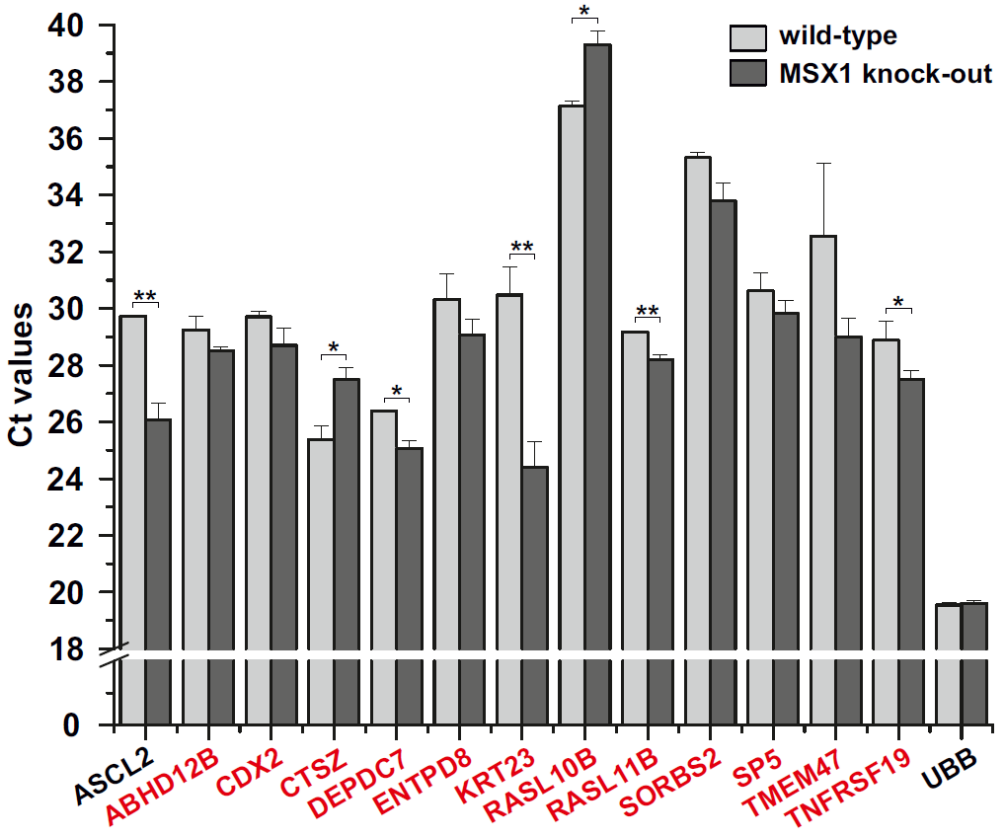


Figure 44 | Analysis of MSX1-deficient SW620 cells.
Quantitative RT-PCR analysis of total RNA samples obtained from SW620 cell clones harboring truncations in the *MSX1* gene (n = 4) and wild-type cell clones (n = 4). The diagram shows Ct values normalized to expression of the internal housekeeping gene *β-actin* (Ct value was arbitrarily set to 17). Genes identified by Watanabe and colleagues[400] are depicted in red. The qRT-PCR reactions were run in technical triplicates; error bars indicate SDs; *, $p < 0.05$, **, $p < 0.01$. *ABHD12B, Abhydrolase Domain Containing 12B; CDX2, Caudal Type Homeobox 2; CTSZ, Cathepsin Z; DEPDC7, DEP Domain Containing 7; ENTPD8, Ectonucleoside Triphosphate Diphosphohydrolase 8; KRT23, Keratin 23; RASL10B, RAS Like Family 10 Member B; RASL11B, RAS Like Family 11 Member B; SORBS2, Sorbin And SH3 Domain Containing 2; SP5, Sp5 Transcription Factor; TMEM47, Transmembrane Protein 47.*

The gene encoding transcription factor ASCL2 displayed robust upregulation in MSX1-deficient cells. In the mouse intestines, Ascl2 synergistically with β-catenin/Tcf4 complexes regulates expression of genes fundamental to the intestinal stem cell identity. To test whether MSX1 directly binds *ASCL2* regulatory regions, we performed a chromatin immunoprecipitation (ChIP) of chromatin isolated from SW620 cells. As commertionally available antibodies were not suitable for MSX1 immunoprecipitation, we used the CRISPR/Cas9-mediated gene targeting to insert the EGFP encoding sequence to the 5' end of the *MSX1* locus (Figure 45) and generated cells endogenously expressing the N-terminally tagged EGFP-MSX1 fusion protein. Anti-EGFP antibodies were used to

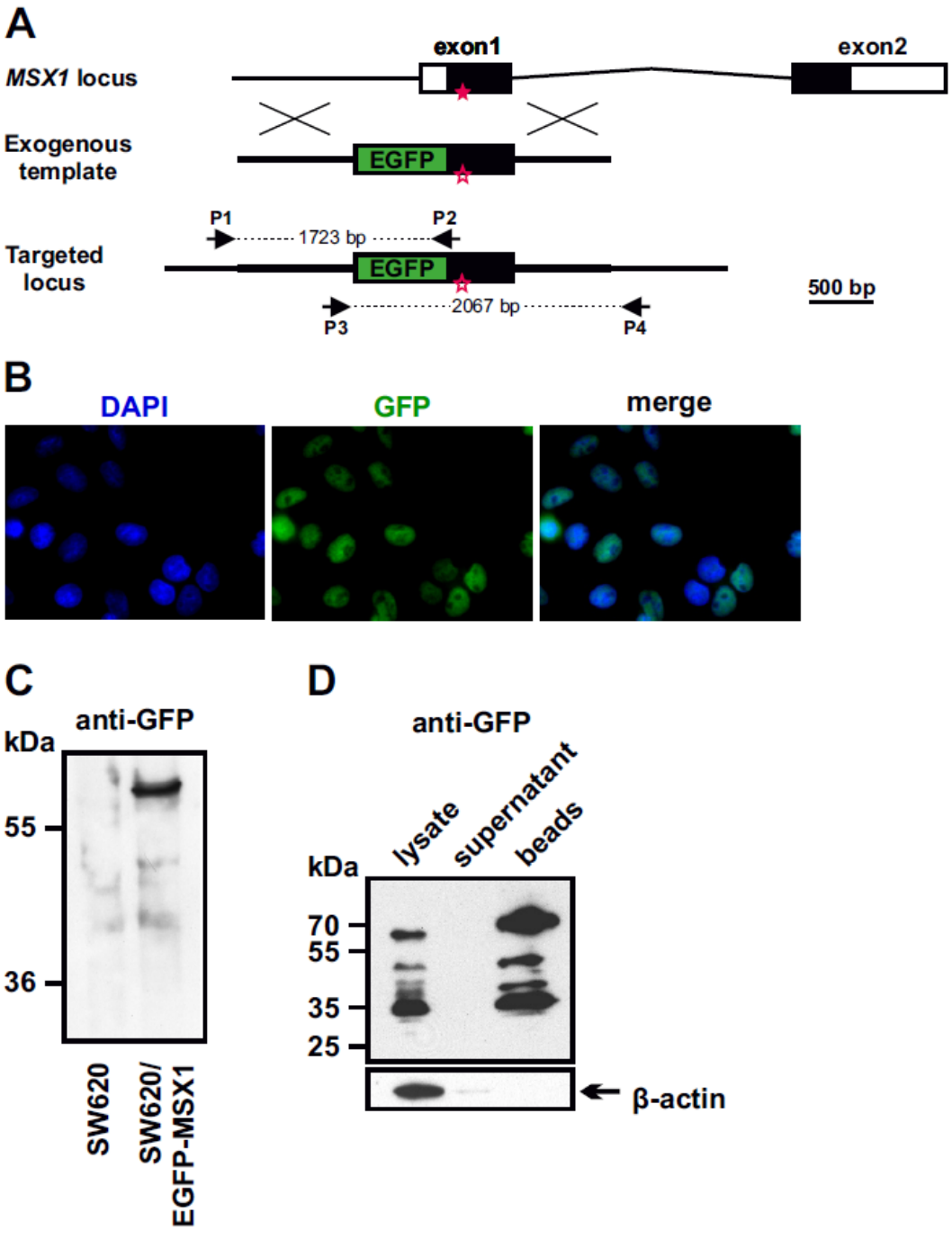


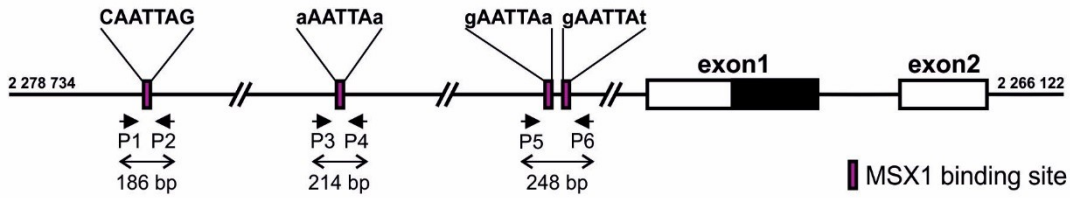
Figure 45 | Generation and analysis of SW620 cells harboring a modified *MSX1* allele producing EGFP-MSX1 fusion protein.
(A) The diagram shows CRISPR/Cas9-based genome editing of the *MSX1* locus; exons are depicted by black boxes, untranslated regions at 5' and 3' end of exon 1 and 2 are indicated as white boxes. CRISPR-Cas9 recognizes and cleaves a specific sequence in the first exon of the gene (magenta asterisk). The affected locus is repaired by homologous recombination using an exogenous template carrying a portion of the *MSX1* gene including the enhanced GFP (EGFP) sequence (green box). CRISPR-Cas9 recognition site in the exogenous template was wobbled (empty magenta asterisk) to prevent its cleavage. Primary screen of clones for presence of the EGFP sequence was done by PCR analysis of genomic DNA using internal EGFP primers originally designed for qRT-PCR (not shown in the diagram). Correct targeting (at both ends of the template) was verified by sequencing of PCR products amplified from genomic DNA using two primer pairs: P1 and P2, P3 and P4. Primer positions are depicted by black arrows; one primer from each set (P1 and P4) primes in a sequence that is not present in the exogenous template. (B) Fluorescent immunocytochemical staining with anti-GFP antibody (green) visualized nuclear localization of endogenously produced EGFP-MSX1 fusion protein in SW620 cells. Cells were counterstained with DAPI nuclear stain (blue). Magnification: 1000 ×. (C) Western blotting of SW620 and SW620/EGFP-MSX1 cell lysates with anti-GFP antibody confirmed endogenous production of EGFP-MSX1 fusion protein. (D) Western blotting with an anti-GFP antibody of SW620/EGFP-MSX1 cell lysate used for immunoprecipitation by GPF-Trap beads, i.e. input (left), supernatant from beads after incubation with the lysate (middle), and precipitate retained on beads (right). Anti-β-actin antibody was used as a loading

immunoprecipitate EGFP-MSX1 fusion protein crosslinked to the sonicated chromatin fragments. Quantitative RT-PCR analysis was performed to analyse changes in levels of DNA fragments containing putative MSX1-binding sites that we identified in the *ASCL2* promoter sequence. However, the analysis did not reveal any MSX1 binding sites in the *ASCL2* regulatory region (Figure 46A, C).

A similar analysis was performed to identify MSX1-binding sites in the regulatory region of the *SP5* gene. Previous publications suggested a presence of β-catenin/TCF4 recognition sites in the *SP5* promoter in human CRC cells[356, 400] and our analysis of the *SP5* promoter identified several putative MSX1 binding sites. However, the ChIP and following qRT-PCR analysis did not indicate MSX1 binding to the *SP5* promoter (Figure 46B,C) and luciferase reporter assay did not reveal MSX1-dependent regulation of *SP5* (Figure 46D).

Subsequently, the regulatory region of the gene encoding *Sox17* was analyzed. The *Sox17* mRNA levels were robustly upregulated in Apc-/Msx1-double-deficient mouse intestinal epithelial cells and detected in the mouse small intestinal tumors (Figure 36). Additionally, previous studies reported that *Sox17* is expressed in the intestinal epithelium and functions antagonistically to the β-catenin/TCF complexes[335]. To test responsiveness of the *Sox17* promoter region to MSX1 transcription factor, a 478 bp long DNA sequence including two MSX1 binding sites in close proximity to the promoter was PCR-amplified

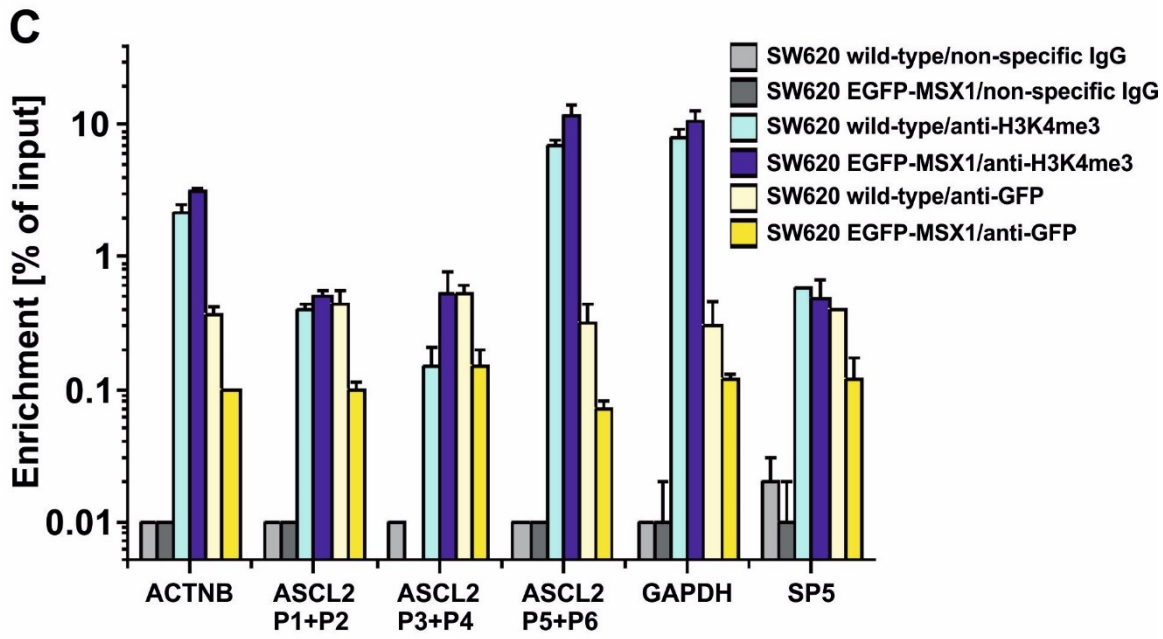
A ***ASCL2* locus**



B ***SP5* promoter region**

GCGGGTACCGCGAGGGTGCAGGGTGTGCAAGTAAATACAGGGGAGAGTTACACTCTTTGCCG
TCTCAAAGTCCGGCGCCTGGAAGCCGAAGGCAGATTTTTCCAGAGATAACAAAGACACTTTG
TCACTTTGGGCGCTGCGTCGGTGCAAATCTTTAATTGCAAGTGTGTGTGTGTGTGTGTGTGT
GTATGTTGGGGGGGGCGGGAAGAGTGTCTCGAGCAACCCAAGCGCGGGTCTCCAGGCGGCAA
GGCCCCCTTTGATCAGGAAAATCCAATTATTTGTGTATA**TACATTTCCCACATAGTGATTAC**
TTCATTAATTGTCCCGCCTTTATCTCCTCCCTTCCCATCCCCCCTAATAATCAGTTCTTTTA
TCCAGACCAACAAACACACCATAGGAGCTTTGTGGATTCAAAGGATTTGCTTTCGCTTCTGA
AAGAGCCGCTATTCTTTGATGATTGGGTAGCGGCAAACTTCAAAGCCATAAATCTTCCCTCT
GACTGGCTGGCGGCCCAGCAAAGTCCTTATCAAATTCTTGGAGGTGATCCTGAAGCGCTCAG
ACCGCGCGCGGGGCGAGCGAGCGGGGCGCGGCGAGGGGCAAGGGCGGGGAGGGCCCCGGCGC
TCAGAGCAGGCGCCAGGGAGGCAGGCTGGGCGGCCCTTCGTCCTCGCCTTCGGGTGTCCATG
CCTCGGCGGCGGCGTCCCGCTCCGCAGCCAGGGGCCTGCAAGCCGTAGCCATGGCCGCGGTG
GCCGTCCTCCGGAACGACTCGCTGCAGGCCTTTCTCCAGGTCAGGGCCGAGCCCGGAGGGGG

C



D

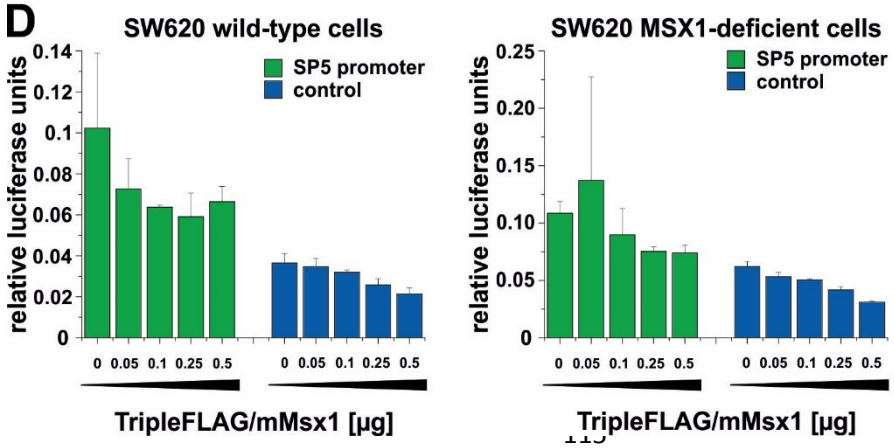


Figure 46 | ChiP analysis of the *ASCL2* and *SP5* regulatory regions.
(A) The diagram shows potential MSX1 binding sites (pink boxes) in the *ASCL2* locus; exons and untranslated regions are depicted by filled or empty boxes, respectively. The sequences of consensus MSX1 binding sites are shown, nucleotides surrounding the MSX1 binding core not matching the consensus sequence are indicated by lowercase letters. The ChIP analysis was performed using indicated primer pairs; primer positions are depicted by black arrows. The numbers above the diagram indicate nucleotide positions in the human genome assembly GRCh38:CM000673.2. (B) The sequence of the *SP5* gene regulatory region with β-catenin binding peak (in bold) that was identified by Watanabe and colleagues by integrative ChIP-seq/microarray analysis in colon cancer cells[400]. TCF/β-catenin and MSX1 binding sites are depicted in magenta and blue, respectively; *MSX1* ChIP primers are indicated by green lines; transcription start site is indicated by a red arrow; initiation codon is highlighted in yellow. The sequence maps to nucleotides 170,714,781-170,715,617 on chromosome 4 in human GRCh38 coordinates. (C) Crosslinked and sonicated chromatin obtained from SW620 wild-type and EGFP-MSX1 fusion protein producing cells was precipitated with following antibodies: non-specific IgG (negative control), H3K4me3 (positive control - recognizes chromatin of transcribed genes), and GFP. ChIP analysis of housekeeping genes *β-actin* (*ACTB*) and *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) using the anti-H3K4me3 antibody was used as a positive control. The recovery of the particular promoter region was calculated as the relative amount of immunoprecipitated DNA compared to input DNA, i.e. percentage of the input. (D) Luciferase reporter assay in SW620 cells harboring the truncated *MSX1* alleles does not indicate *SP5* gene regulation by MSX1. Relative luciferase units (RLU) indicate level of luciferase activity normalized to the number of transfected cells (measured by activity of renilla enzyme). A luciferase reporter vector containing the *SP5* promoter sequence with three consensus MSX1 binding sites (nAATTAn; n stands for any nucleotide) was analysed; an empty luciferase reporter vector was used as a control. Samples were measured in technical duplicates; experiment was performed in two replicates, representative results are shown.

from the *Sox17* locus and cloned into the luciferase reporter vector pGL4.26 (containing a minimal promoter); alternatively, a 30 bp long dsDNA fragment with the two MSX1 binding sites was cloned into the pGL4.26 vector. Luciferase-reporter assay was performed in SW620 wild-type cells and cells harboring truncations in the *MSX1* gene. However, the assay did not reveal MSX1-dependent regulation of the tested DNA sequences (Figure 47A, B). Previously published studies also reported that *Sox17* expression is reduced in $Apc^{+/Min}$ mouse intestinal tumors and during adenoma progression to carcinoma[254]. To determine *SOX17* mRNA levels in human tumors, qRT-PCR analysis in a collection of RNA samples isolated from human colorectal neoplasia specimens and matching healthy mucosa was performed. *SOX17* displayed slight decrease in early hyperplastic adenomas and increasing tendency with progression to dysplasia and carcinoma (Figure 48A). *SOX17* expression was then analyzed in human colorectal tumors with tubular, tubulovillous or villous morphology. However, no correlation between tumor morphology and *SOX17* mRNA

levels was observed (Figure 48B). Finally, a correlation analysis of *MSX1* and *SOX17* was performed on the collection of RNA samples isolated from human colorectal neoplasia. Interestingly, a positive correlation between the mRNA levels of *MSX1* and *SOX17* was observed, although it was not statistically significant (Figure 48C).

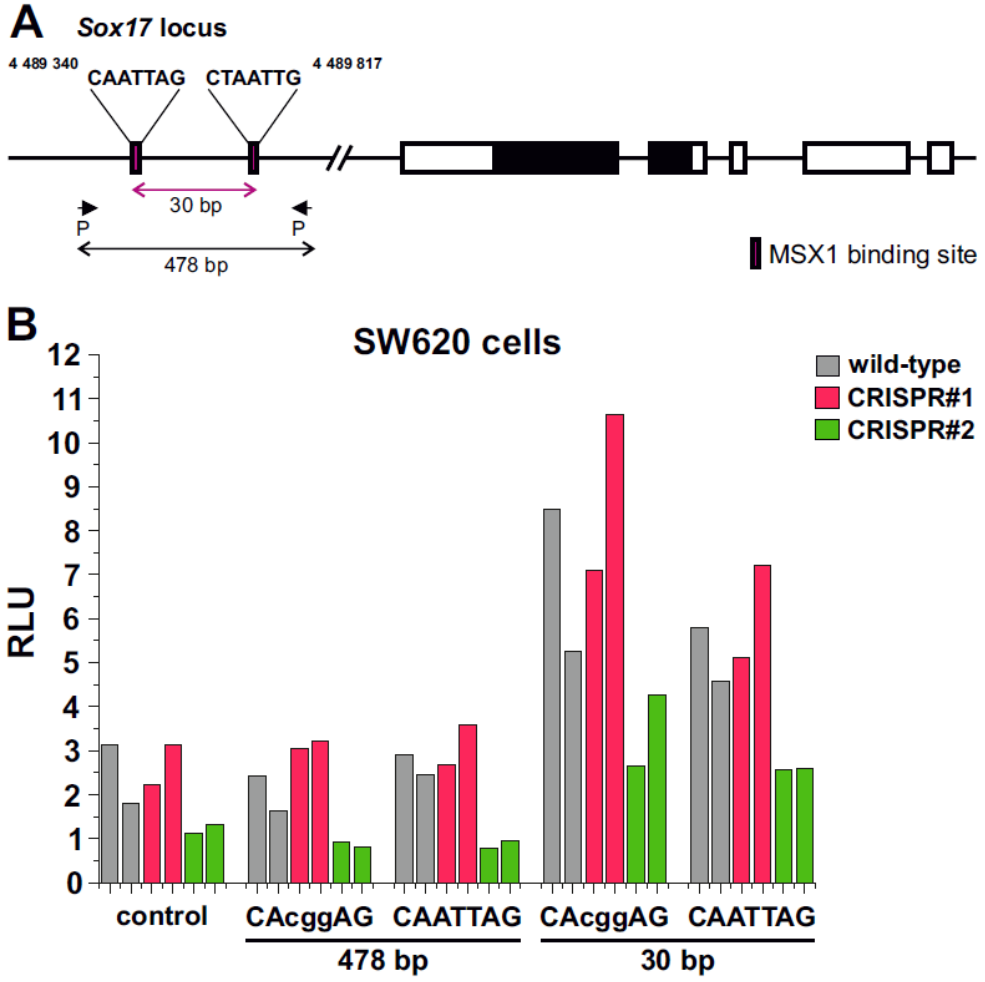


Figure 47 | Analysis of the *Sox17* regulatory region. (A) The diagram shows MSX1 binding sites (pink boxes) in the *Sox17* locus; exons and untranslated regions are depicted by filled or empty boxes, respectively. The sequences of consensus MSX1 binding sites are shown; notice that the sites are in a palindromic orientation. The 478 bp sequence indicated by black double arrow was PCR amplified and cloned into the pGL4.26 luciferase reporter vector. Similarly, the 30 bp sequence containing two MSX1 binding sites (pink double arrow) was cloned into the pGL4.26 plasmid as annealed complementary oligos. The numbers above the diagram indicate nucleotide positions in the mouse genome assembly GRCm38:CM000994.2. (B) Luciferase reporter assay in SW620 cells harboring the truncated *MSX1* alleles does not indicate *Sox17* gene regulation by MSX1. Relative luciferase units (RLU) indicate level of luciferase activity normalized to the number of transfected cells (measured by activity of renilla enzyme). Two luciferase reporter vectors containing consensus MSX1 binding sites (CAATTAG) within 478 or 30 bp long inserted DNA sequences were analysed. In addiotion, two reporters with mutated MSX1 sites (CAcggAG; altered nucleotides are indicated in lowercase letters) were tested. Empty luciferase reporter vector was used as a control. Samples were measured in technical duplicates; experiment was performed in two replicates, representative results are shown.

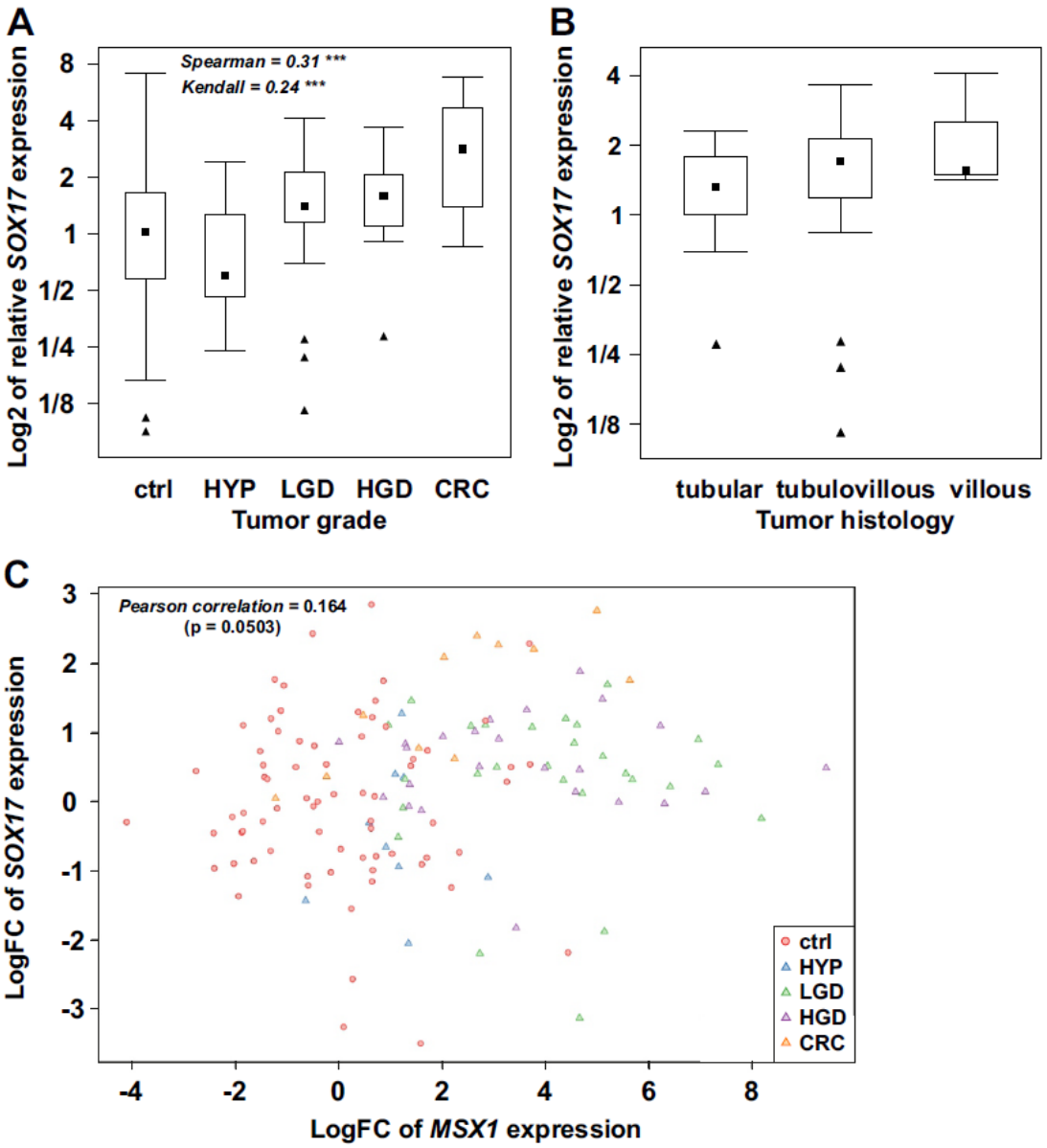


Figure 48 | *SOX17* expression increases during colorectal neoplasia progression but is not directly linked to the histological characteristic of human adenomas. (A, B) Quantitative RT-PCR analysis of *SOX17* mRNA in samples obtained from human colorectal neoplasia specimens. The samples were divided into categories based on the tumor grade (A) or tumor histology (B) as follows: (A) healthy tissue (ctrl), hyperplastic adenomas (HYP; n = 9), adenomas displaying low-grade dysplasia (LGD; n = 27) or high-grade dysplasia (HGD; n = 24), and colorectal carcinoma (CRC; n = 12); (B) tubular (n = 20), tubulovillous (n = 31), and villous (n = 3). The amounts of RNA in individual isolates were normalized to the geometric average of Ct values of housekeeping genes *UBB* and *β2-microglobulin*. The boxed areas correspond to the second and third quartiles; the median of ΔCt values for each group of samples is indicated as the black square. The range of the values is given by "whiskers" above and below each box; outliers are indicated by black triangles. The significance between the *SOX17* mRNA level and neoplasia progression is demonstrated by the Spearman ($\rho = 0.31$) and Kendall ($\tau = 0.24$) coefficient values; ***, $p < 0.001$. (C) Correlation analysis of *SOX17* and *MSX1* expression in human colorectal neoplasia. The positive correlation is demonstrated by Pearson correlation coefficient ($\rho = 0.164$), however, it was not statistically significant ($p = 0.0503$).

4.4 HIC function in mouse intestines

The following chapter describes the results of a project that focused on investigation of the Hic1 tumor suppressor in the mouse intestine. Results of experiments I participated in are shown. For a better understanding of the topic, several experiments I did not participate in are mentioned in the text. Such results are, however, not shown in figures of this thesis, and may be found in the article by Janeckova and colleagues[146].

Hic1-deficient small intestinal epithelium contains increased numbers of Paneth, goblet, and enteroendocrine cells

To examine function of the Hic1 tumor suppressor, Pospichalova and co-workers generated a mouse strain harboring conditional alleles of the *Hic1* gene[271]. The *Hic1*$^{cKO/cKO}$ mice were intercrossed with *ROSA-CreERT2* animals expressing inducible CreERT2 recombinase in all cell types. Mouse embryonic fibroblasts (MEFs) were isolated from *Hic1*$^{cKO/cKO}$ *ROSA-CreERT2* mice and *Hic1* gene inactivation was induced by administration of 4-OHT; control cells were administered with the solvent (ethanol) only. The MEFs were harvested at several timepoints (24, 48, 72, and 120 hours) after 4-OHT administration, total RNA was isolated and microarray analysis was performed. The gene expression profiling revealed a subset of differentially expressed genes (with significance criteria $|\log FC| \geq 1$ and q-value < 0.05) in at least two timepoints, including toll-like receptor 2 (*Tlr2*). As TLR-mediated signaling has been linked to inflammation-associated tumorigenesis in the colon and rectum[336], the relationship between *TLR2* and *HIC1* was analyzed in more detail. Luciferase reporter assay and ChIP analysis revealed direct regulation of the *TLR2* gene expression by HIC1 and siRNA-mediated downregulation of *HIC1* resulted in the increased *TLR2* levels in human cells, confirming that *TLR2* is the HIC1 target gene.

Hic1$^{cKO/cKO}$ mice were further intercrossed with *Villin-Cre* mice expressing constitutively active Cre recombinase in all intestinal epithelial cells. As the Hic1-deficient intestines lacked any (observable) morphological changes and epithelial cells did not exhibit impaired proliferation, we performed a detailed analysis of individual populations of differentiated cell lineages. Small intestinal crypt cells obtained from *Hic1*$^{cKO/cKO}$ *Villin-Cre* (*Hic1*$^{KO/KO}$) and control (*Hic1*$^{+/+}$) mice were stained with anti-EpCAM (marker of epithelial cells), anti-CD45 (marker of leukocytes), and anti-CD24 (marker of the crypt base cells) antibodies and a subpopulation of $CD45^{-}EpCAM^{+}CD24^{+}$ cells (i.e. Paneth cells)

was obtained. The FACS-analysis showed slightly increased amount of Paneth cells in the Hic1-deficient crypts compared to Hic1-proficient crypts (Figure 49A). Paneth cells maturation is under control of the transcription factors Atoh1[330] and Sox9[227], which have been described as Hic1-target genes in the developing mouse cerebellum[31] and in human osteosarcoma cells[384], respectively. Quantitative RT-PCR analysis confirmed, besides decreased expression of *Hic1*, elevated levels of *Atoh1* and *Sox9* mRNA and, moreover, also increased expression of *Tlr2* in Hic1-deficient Paneth cells (Figure 49 B). In addition, immunohistochemical staining of the small intestine was performed using an anti-lysozyme antibody, which specifically stains Paneth-cells. Subsequent analysis confirmed increased number of Paneth cells in the $Hic1^{KO/KO}$ small intestine and, moreover, revealed that the amount of Paneth cells increases along the rostro-caudal axis towards the distal part of the small intestine, regardless of the Hic1 absence (Figure 49 C). Moreover, (immuno)histochemical staining of goblet and enteroendocrine cells by Periodic acid-Schiff (PAS) or an anti-chromogranin A antibody, respectively, revealed increased numbers of both cell types in the Hic1-deficient small intestine (Figure 50A, B). In contrast, number of absorptive enterocytes was unchanged, as evidenced by qRT-PCR analysis of enterocyte-specific markers hairy and enhancer of split-1 (*Hes1*) and sucrase isomaltase (*SI*) and by histological staining of alkaline phosphatase produced by differentiated enterocytes (Figure 50C).

Study of Mohammad and co-workers described accelerated formation of intestinal polyps developed in $Apc^{+/Min}$ mice upon loss of a single *Hic1* allele[223]. Since inflammation is a promoter of intestinal tumor development and the Hic1 target gene *Tlr2* is an important component of the inflammatory immune response in the intestine (reviewed in[162]), the function of Hic1 was examined also in the mouse model of colitis-associated tumorigenesis. Acute colitis was induced in $Hic1^{cKO/cKO}$ *Villin-Cre* ($Hic1^{KO/KO}$) and control ($Hic1^{+/+}$) mice by DSS addition to drinking water; nevertheless, $Hic1^{KO/KO}$ colon did not exhibit any difference in the tissue damage in comparison to the $Hic1^{+/+}$ colon. However, during the regenerative phase, i.e. 6 and 9 days upon DSS withdrawal, the $Hic1^{KO/KO}$ colonic epithelium exhibited more robust proliferation. Finally, combined treatment with AOM and DSS resulted in significantly increased size of tumors in the $Hic1^{KO/KO}$ colon. Therefore we suppose that in the intestines, Hic1 functions as a tumor suppressor and its depletion from the tissue partially deflects the balance of cell differentiation towards the secretory lineages. More importantly, the Hic1 transcription factor is particularly important during regeneration upon inflammation-induced damage of the tissue.

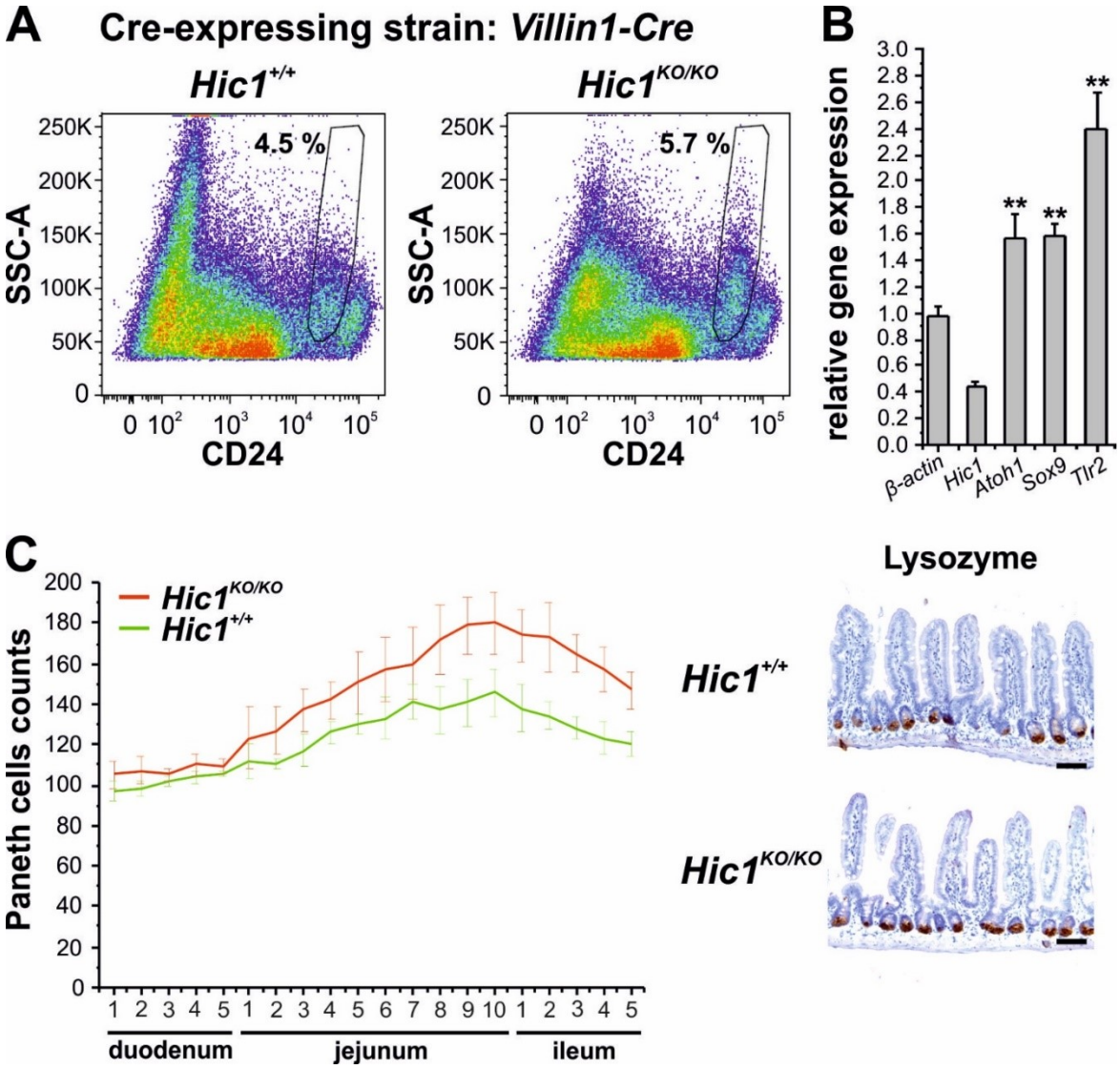


Figure 49 | Hic1 ablation from the mouse small intestinal epithelium results in an increased number of Paneth cells. (A) Small intestinal crypt cells obtained from *Hic1*$^{cKO/cKO}$ *Villin-Cre* (*Hic1*$^{KO/KO}$) and control (*Hic1*$^{+/+}$) mice were stained and FACS-sorted to obtain CD45^{-} EpCAM^{+}CD24^{+} population of Paneth cells. The histograms show increased number of Paneth cells in *Hic1*$^{KO/KO}$ tissue. Four animals from both strains were used; the experiment was performed twice; representative images are shown. (B) Quantitative RT-PCR analysis of RNA obtained from sorted Paneth cells. The diagram shows decreased levels of *Hic1* mRNA and elevated expression of Hic1 target genes *Atoh1*, *Sox9*, and *Tlr2* in *Hic1*$^{KO/KO}$ cells relative to *Hic1*$^{+/+}$ cells (gene expression levels in control cells were arbitrarily set to 1). The gene expression levels were normalized to internal housekeeping gene *UBB*; PCR reactions were run in triplicates; the experiment was performed twice; representative results are shown. Error bars indicate SDs; **, $p < 0.01$. (C) The diagram shows distribution of Paneth cells in the indicated segments of the small intestine. Sections obtained from *Hic1*$^{KO/KO}$ and *Hic1*$^{+/+}$ mice were immunohistochemically stained by anti-Lysozyme antibody (brown signal) to visualize Paneth cells (right). Lysozyme-positive cells were counted in 50 neighboring crypts in several different fields (duodenum: n = 5, jejunum: n = 10, ileum: n = 5). Four animals from both strains were used in the experiment; representative images are shown. The sections were counterstained with hematoxylin; error bars indicate SDs; scale bar: 0.15 mm.

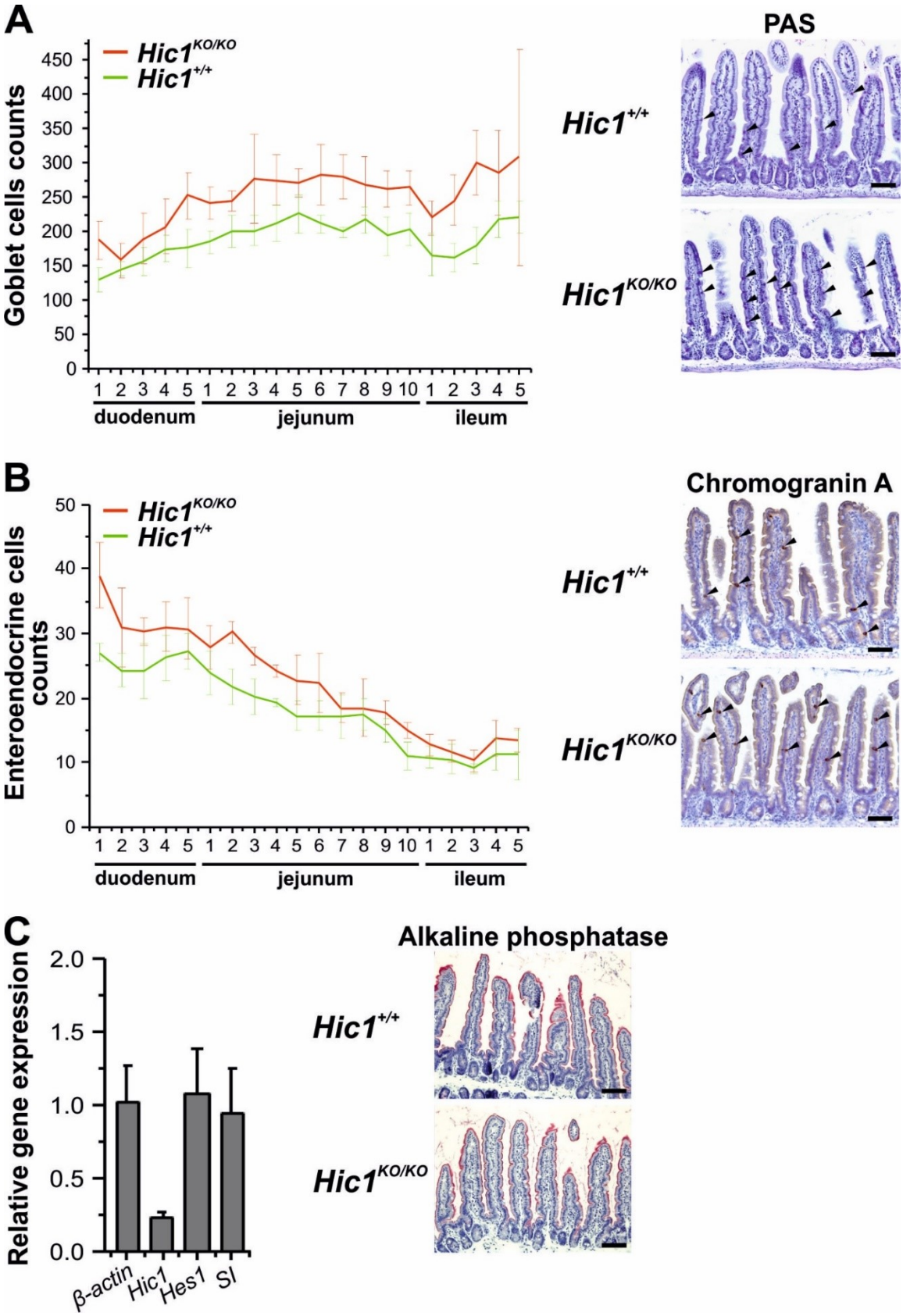

A
Goblet cells counts
Hic1KO/KO
Hic1+/+
duodenum
jejunum
ileum
PAS
Hic1+/+
Hic1KO/KO
B
Enteroendocrine cells counts
Chromogranin A
C
Relative gene expression
β-actin
Hic1
Hes1
SI
Alkaline phosphatase

Figure 50 | Hic1 ablation from the mouse small intestinal epithelium results in increased number of goblet and enteroendocrine cells.
(A, B) The diagrams show distribution of goblet (A) and enteroendocrine (B) cells in the indicated segments of the small intestine. Sections obtained from *Hic1*$^{KO/KO}$ and *Hic1*$^{+/+}$ mice were stained using Periodic acid-Schiff (PAS; magenta signal) and an anti-chromogranin A antibody (brown signal) to visualize goblet and enteroendocrine cells, respectively (indicated by black arrowheads). Cells were counted in 50 neighboring crypts in several different fields (duodenum: n = 5, jejunum: n = 10, ileum: n = 5). Four animals from both strains were used in the experiment; representative images are shown. The sections were counterstained with hematoxylin; error bars indicate SDs; scale bar: 0.15 mm. (C) Left, qRT-PCR analysis of RNA obtained from *Hic1*$^{KO/KO}$ and *Hic1*$^{+/+}$ small intestinal epithelial cells. The diagram shows expression of enterocytes-specific markers hairy and enhancer of split-1 (*Hes1*) and sucrase isomaltase (*SI*) in *Hic1*$^{KO/KO}$ cells normalized to internal housekeeping gene *UBB* and relative to *Hic1*$^{+/+}$ cells. Gene expression levels in control cells were arbitrarily set to 1; another housekeeping gene *β-actin* is shown; error bars indicate SDs. Right, staining of alkaline phosphatase (magenta signal) produced by differentiated enterocytes. The sections were counterstained with hematoxylin; scale bar: 0.15 mm.

5 Conclusions

Although the function of the Msx1 protein has been thoroughly described in the mouse embryonic development, due to the perinatal lethality of the *Msx1*$^{-/-}$ mice, there were not many studies that would address the Msx1 function in adult tissues. Additionally, the role of Msx1 in the intestines has not been described at all. In this thesis, we aimed to clarify the function of Msx1in the healthy intestine and its involvement in development of intestinal tumors. Our aim was to identify the effect of *Msx1* gene inactivation on intestinal morphology and tumor growth. We observed that under homeostatic conditions, *Msx1* expression in the mouse intestines was virtually undetectable; however, after inactivation of the gene encoding the tumor suppressor Apc, *Msx1* expression significantly increased. Although we failed to describe the Msx1 function in intestinal tumorigenesis at the molecular level, our observations clearly show that Msx1 is essential in the process of ectopic crypt formation. In addition, it is very likely that Msx1 plays an important role in development of intestinal tumors, as its inactivation causes morphological conversion towards villous adenomas, which represent a more advanced type of CRC. Concordantly, in human intestinal tumors *MSX1* is most abundant in early stages of neoplasia with a decreasing tendency towards carcinoma. Nevertheless, studies published previously that focused on the MSX1 role in human cancer often brought contradictory conclusions. Therefore, another objective of this work was to describe the MSX1 function in human intestinal tumors. Unfortunately, our results did not provide a clear answer to the question of how MSX1 affects human tumorigenesis. *MSX1* inactivation in human CRC cells did not affect cellular proliferation or migration. Moreover, in MSX1-deficient CRC cells we did not observe any differences in their ability to xenograft when injected into the immunodeficient mice. However, our results indicate that *MSX1* represents a robust marker of human colorectal tumors.

Although we utilized different approaches to identify MSX1-regulated genes in the mouse intestinal epithelium and human tumors, we have not succeeded in identification of MSX1 binding sites in the regulatory regions of selected genes, whose expression changed upon MSX1 depletion. However, the gene expression profiling in human CRC cells revealed a significant overlap between putative MSX1-regulated genes and a group of genes, whose regulatory regions were bound by β-catenin. We therefore suggest that MSX1 is involved in transcriptional regulation of a specific gene subset that is regulated by other effector proteins of the Wnt signaling pathway as well.

Finally, the last aim of this thesis was to investigate the effect of inactivation of the tumor suppressor gene *Hic1* in mouse intestinal tissue. The results presented in this theses represent only a part of the Hic1 project. Specifically, I contributed to the Hic1 "story" by analyzing changes in the representation of various differentiated cell types present in the small intestinal epithelium. We confirmed that Hic1 downregulates expression of *Atoh1* and *Sox9*, two genes encoding transcription factors which have important functions in differentiation of intestinal epithelial cells. Depletion of Hic1 resulted in upregulation of these two genes and subsequent increase in numbers of Paneth, goblet, and enteroendocrine cells. Further results of the Hic1 project, which are not included in this theses, suggest that Hic1 downregulates expression of *Tlr2*. Inasmuch as the Tlr2-mediated activation of the NF-κB signaling pathway promoted intestinal tumorigenesis, we suggest that *Hic1* inactivation might potentiate tumor-promoting pro-proliferative Tlr2/NF-kB signaling.

6 Discussion

In this study, we aimed to describe the function of two transcription factors, MSX1 and HIC1, which are both linked to the Wnt signaling pathway and (presumably) have tumor suppressor functions. The main part of the study dealt with the role of MSX1 in the mouse intestines and human CRC, while the final chapter of the results described the effect of Hic1 loss in the mouse small intestine.

The initial experiment that led us to study Msx1 in the intestines was based on identification and characterization of the genetic program in transformed epithelial cells after loss of the tumor suppressor Apc, i.e. upon hyperactivation of the Wnt signaling pathway. We used experimental mice harboring conditional alleles of the *Apc* gene and expressing Cre recombinase in all intestinal epithelial cells. The gene expression profiling of cells obtained from the intestines before and after *Apc* inactivation revealed significant changes in expression of many genes, including several previously described Wnt target genes as well as many "new" genes whose function has not been linked to the Wnt signaling pathway in the context of the gastrointestinal tract yet.

One of the genes with the most increased expression after *Apc* inactivation was Msx1 (also known as Hox7), nuclear protein that belongs to the muscle segment homeobox family of transcription factors (Figure 15). In mice, Msx1 has been studied predominantly during embryogenesis where it is involved in numerous processes such as development of the teeth, limbs, and brain[9, 184, 298]. In human, MSX1 has been studied especially in the context of craniofacial development, as mutations in *MSX1* are linked to tooth agenesis, cleft lip and cleft palate[151, 154, 239, 386]. Although the MSX1 function in human colorectal cancer has not been properly described yet, few studies identified *MSX1* promoter hypermethylation in colorectal neoplasia[320, 350, 396], suggesting functional consequences of *MSX1* downregulation in tumor cells[396].

Nevertheless, many other independent research groups presented data more consistent with our results which indicate that Msx1 is a marker of tumor tissue[222, 322, 350]. Moreover, data provided by the BioGPS portal, which allows extensive analysis of data acquired from gene expression experiments, showed that *MSX1* is predominantly expressed in samples obtained from colorectal adenocarcinoma (Figure 38). Moreover, according to data available in the COSMIC database[75], *MSX1* is only rarely mutated in human colon tumors (15 mutated samples in 1564 tested samples). Therefore it is tempting to speculate that in colorectal cancer, *MSX1* works as a tumor-promoting gene and might have some important

functions. As we observed the highest *MSX1* expression in early stages of human intestinal neoplasia with a decreasing tendency towards carcinoma, we suppose that MSX1 is probably important during tumor initiation and in the early stages of CRC development rather than in more advanced carcinoma.

It is generally assumed that Apc depletion from the intestinal epithelium results in hyperactivation of the Wnt signaling pathway, which leads to perturbation of the gene expression program essential for *stemness* of epithelial cells; this is closely associated with deregulated cell proliferation, migration, differentiation, and apoptosis. The Apc-deficient intestinal epithelium exhibits altered crypt-villus architecture with the prolonged crypt compartment containing highly proliferating and abnormally dense packed cells; moreover, morphologically atypical crypt-like cells occupy majority of the villous compartment adjacent to the crypts[305]. Although it is believed that the sole inactivation of the *Apc* gene does not lead to ectopic crypts formation[16], we observed fully developed ectopic crypts in the villi seven days upon *Apc* inactivation (Figure 20). Moreover, our data show that in the Apc-deficient small intestinal epithelium, the seemingly continual hyperplastic crypt compartment consists of two parts, the "standard" ISC compartment with crypts oriented to the underlying mucosa, and the proliferating region in the villi containing the Msx1-positive ectopic crypts. The organization of the intestinal epithelium is driven, apart from ISCs division, by generation of new crypts by fission[345]. This program, which is highly active in developing embryos and attenuated postnatally, is restored in cancer tissue and promotes tumorigenesis[6, 132, 279]. Moreover, under some (pathological) circumstances, abnormally oriented crypts (i.e. ectopic crypts) located in the villi may develop. Several independent studies described possible ways of the ectopic crypts formation in the mouse intestine. Aberrant expression of the BMP signaling pathway inhibitor noggin or gremlin 1 initiated budding of intravillous ectopic crypts followed by development of dysplastic features that progressed to polyposis[17, 49, 102]. Madison and colleagues observed that inhibition of the Hedgehog signaling pathway impaired the crypt-villus architecture, which was accompanied by ectopic crypts formation[203]. In addition, aberrant expression of the Ascl2 transcription factor induced ectopic crypts development accompanied by increased proliferation of the crypt compartment[377]. Proliferating aberrant foci were also observed upon simultaneous hyperactivation of the NF-κB and Wnt signaling pathways. As these foci expressed the stem cell markers *Ascl2* and *Lgr5*, the authors proposed that formation of the proliferating aberrant foci was triggered in villus cells by induction of a dedifferentiation program[329]. Recently, Perekatt and colleagues also reported that ectopic

crypts may arise from differentiated cells. Interestingly, inactivation of the differentiation-promoting transcription factor SMAD4 and simultaneous activation of the Wnt signaling pathway caused that enterocytes re-entered the cell cycle, restored expression of stem cell markers, and initiated formation of ectopic crypts[265]. We cannot rule out the possibility that the ectopic crypts are generated by cell dedifferentiation also in our experiments. However, our observations of proliferating and Msx1-positive cells at various time points after *Apc* inactivation rather suggest that the ectopic crypts arise from cells that originated from the "standard" crypt compartment and which are aberrantly located in the villi (Figure 20 and 27). In human, the presence of ectopic crypts represents a characteristic feature of serrated adenomas, a minor fraction of CRC which typically progress to aggressive tumor stages[22]. Although we have observed that Msx1 marks the ectopic crypts in mice, our analysis of human CRC revealed elevated *MSX1* expression across all stages of colorectal tumors (Figure 39A), indicating an important role of MSX1 in cancer initiation and progression.

While *Msx1* activation in the intestinal tissue is closely associated with Apc loss, Msx1-positive cells are not present in the hyperplastic crypt compartment, but are located in the villi. This indicates that another regulatory mechanism is probably involved in *Msx1* gene expression in the intestines. Wallmen and colleagues discovered that the transcription factors from the TCF/LEF family are not able to bind silent chromatin and initiate *de novo* expression of inactive genes, thus ensuring spatiotemporal stability of the Wnt target genes expression[392]. Interestingly, numerous independent studies reported that during embryonic development and in human cancer cells, *Msx1* expression is regulated by the BMP signaling pathway[214, 216]. In the intestines, the activity of BMP signaling is restricted to the villous compartment and decreases towards the crypts, where it is locally attenuated by production of BMP antagonists[103, 175]. Some BMP-activated SMAD transcription factors are (in association with other proteins) able to induce specific DNA demethylation and rearrangement of repressive histone marks, thus allowing initiation of gene expression (reviewed in[210]). We propose that once the recombined cells migrate out of the crypt compartment, the chromatin within the *Msx1* locus is "opened" by protein complexes containing SMAD transcription factors and subsequently the β-catenin/Tcf complexes can activate *Msx1* expression.

Although our results strongly suggest that Msx1 is indispensable for ectopic crypt formation after *Apc* inactivation, the molecular mechanism of how Msx1 influences formation of the ectopic crypts is not clear. The Apc-/Msx1-double-deficient small intestinal epithelium does not form the ectopic crypts, but remains smooth. This is

accompanied by increased number of proliferating PCNA-positive cells that reach up to the top of the villi. In contrast, the number of differentiated cells is reduced, as evidenced by significantly decreased expression of the differentiated cells marker H3K27me3 (Figure 28). Immunohistochemical staining of Paneth and goblet cells also showed a slight decrease in Apc-/Msx1-double-deficient intestinal epithelium; nevertheless, this observation was not confirmed at the level of mRNA expression of respective markers. Paneth cells marker lysozyme was not reduced in the Msx1-deficient epithelium at all, and reduction of another markers cryptdins was not statistically significant. In contrast, goblet cells marker mucin 2 exhibited, surprisingly, slightly increased expression in the Msx1-deficient intestinal epithelium (Figure 29). Additionally, the villi in the Apc-/Msx1-double-deficient epithelium contained large amounts of cell clusters positive for the stem cell markers *Ascl2* and *Olfm4* (Figure 22), suggesting that the expression program of these cells is altered. Nevertheless, these cell clusters are not able to create the intravillous invaginations and establish the ectopic crypts. We suppose that the path leading to the ectopic crypts formation is initiated but for an unknown reason the ectopic crypts fail to develop. Our observations thus implicate that formation of the ectopic crypts is interconnected with intestinal epithelial cells differentiation and that Msx1 inactivation impairs the process.

We further investigated the consequences of *Msx1* depletion from the intestinal epithelium on the level of cell cycle regulation. Several studies described that MSX1 may influence cell proliferation by transcriptional regulation of genes responsible for cell cycle progression[257, 422]. Although we observed increased percentage of cells in the G2 and S phase in the Apc-deficient epithelium, there was no difference between Apc-deficient and Apc-/Msx1-double deficient epithelium (Figure 30). We suppose that Msx1-dependent regulation of the cell cycle might be cell type specific. Alternatively, it is possible that the epithelial cells have already reached their maximal proliferation capacity after *Apc* depletion and further increase after *Msx1* inactivation is not possible. In addition, we observed that Msx1 depletion from the *Lgr5-EGFP-IRES-CreERT2* intestines caused histological changes of developed adenomas reminding conversion from tubular to villous morphology (Figure 32). In humans, the villous adenomas are often larger than the tubular or tubulovillous adenoma and display more severe degree of dysplasia and a higher risk of malignant transformation. Nevertheless, in our collection of human CRC samples, we did not observe any correlation between the adenoma morphology and *MSX1* expression (Figure 39B). Since Msx1-deficient tumors exhibited features typical for more advanced stages of CRC, we postulated a hypothesis that Msx1-deficient mice could have a worse

survival rate in comparison to mice with wild-type *Msx1*. We therefore performed analysis where we compared survival of mice after *Apc* inactivation or simultaneous inactivation of *Apc* and *Msx1*. Surprisingly, we observed slightly prolonged lifespan of the Apc-/Msx1-double-deficient mice; however, statistical analysis showed that this difference was not significant (Figure 31).

As the whole-body inactivation of the *Msx1* gene in mice results in perinatal lethality[311], the impact of Msx1 loss on the adult intestines has not been described. Besides, studies employing conditional *Msx1* alleles did not deal with the intestinal tissue. In our experiments, conditional inactivation of *Msx1* in the adult intestinal epithelium had no apparent impact on the intestinal morphology or histology. The absence of any visible phenotype may be due to the fact that Msx1 loss is compensated in intestinal tissue by the *Msx2* gene. The mouse family of *Msx* genes has three members: *Msx1*, *Msx2*, and *Msx3*[48]. While *Msx3* is expressed only in the dorsal neural tube[323, 399], *Msx1* and *Msx2* expression is much broader and often overlapping. Both genes have important functions in embryogenesis, e.g. in development of limbs, teeth, craniofacial bones, and heart[111, 127, 128, 184, 288]. Interestingly, *Msx1* and *Msx2* often exhibit functional redundancy[127, 128, 184], but in some embryonic tissues their expression patterns are complementary[199]. *Msx2*$^{-/-}$ mice are viable, although they have numerous developmental defects in the skin, teeth, mammary glands, cartilages, and bones[140, 312]. *Msx1*$^{-/-}$ *Msx2*$^{-/-}$ mice display more severe phenotype than the single mutants and die prenatally[9, 100, 141, 184]. As these two transcription factors are closely functionally redundant, we plan to inactivate both genes in mouse intestines to see whether the phenotype we observed in the Msx1-deficient tissue is somehow biased by the activity of Msx2. In addition, *MSX2* has been described as the Wnt target gene[426]. In our experiments, we usually observed almost identical trends in *MSX1* and *MSX2* expression, especially in human cells. Moreover, *MSX2* expression was elevated in human tumors and positively correlated with *MSX1* expression. Overall, it seems likely that *MSX1* and *MSX2* genes are redundant and regulated by similar mechanisms.

In mice, the *Msx1* expression levels are very low in the healthy intestinal epithelium (Figure 19), therefore the absence of any visible phenotype after *Msx1* inactivation could be explained by the fact that under homeostatic conditions, Msx1 has no essential functions in the tissue. To test the possibility that Msx1 does have some functions in the intestinal epithelium under homeostatic conditions, we established organoid cultures from intestinal crypts and monitored their growth and survival after passage. Nevertheless, the Msx1-deficient organoids looked same as the control organoids and displayed no defects in

growth (Figure 25). Moreover, we employed two models of intestinal tissue damage in mice. The mice were either irradiated by a sublethal dose of X-rays which leads to depletion of the stem cell compartment[274], or administered with DSS in drinking water, which induces acute colitis[121]. Nevertheless, in both cases we did not notice any difference in the recovery of the colonic epithelium between control and Msx1-deficient animals (Figure 26).

Analysis of the gene expression program in the $Msx1^+$ ectopic crypt cells revealed increased expression of Wnt target genes *Axin2*, *Lgr5*, and *SP5*. In contrast, the Wnt target gene and stem cell marker *Olfm4* was downregulated, which indicates that the gene signature of ectopic crypt cells differs from the gene expression program of ISCs (Figure 22). Interestingly, while van der Flier and co-workers induced ectopic crypts formation in *Ascl2*-overexpressing transgenic mice[377], we observed increased expression of *Ascl2* in ectopic crypts upon *Apc* inactivation. Nevertheless, ChIP analysis did not confirm direct binding of Msx1 to the regulatory regions of the *Ascl2* gene. We are aware that the FACS-sorting of ectopic crypt cells needs optimization to obtain a clearly separate population of Msx1-positive cells; this should be taken into account when drawing conclusions from these results. A better approach would be to generate transgenic mice harboring an allele expressing fluorescently tagged Msxl protein and sort cells directly based on the endogenous fluorescence. Nevertheless, this approach is very time consuming and at the time of writing this theses, we did not have such transgenic mice.

Alternatively, we attempted to identify genes regulated by Msx1 in Apc-deficient epithelium. However, our analysis did not provide clear results. The significance criteria ($|\log FC| \geq 1$ and q-value ≤ 0.05) were only reached in case of the *Stk32b* gene. We are aware that *Msx1* expression in Apc-/Msx1-double-deficient epithelium was slightly increased in comparison to control epithelium. This is not surprising since the expression of *Msx1* in the healthy intestinal epithelium is undetectable (normalized Ct values > 42). In addition, immunohistochemical staining upon simultaneous inactivation of *Apc* and *Msx1* revealed residual production of the Msx1 protein in cells that apparently escaped recombination (Figure 27A). Anyway, *Msx1* expression in the Apc-/Msx1-double-deficient epithelium is still significantly reduced in comparison to Msx1-proficient epithelium (ΔCt values > 4), which indicates that incomplete *Msx1* inactivation should not impair the results of the gene expression profiling. We suppose that the gene expression profiling of the entire epithelium did not have sufficient resolution to detect all differences between Msx1-deficient and Msx1-proficient epithelium. The *STK32b* gene encodes a serine-threonine

protein kinase and its deletion has been associated with Ellis-van Creveld syndrome, a human disease affecting development of bones, teeth, and heart. *STK32b* was also described as a putative biomarker of more aggressive types of breast tumors and its elevated expression was associated with smaller size of oral squamous cell carcinoma[259, 364]. Moreover, single-nucleotide polymorphisms (SNPs) in the *STK32b* gene have been described in patients with non-syndromic orofacial clefts[137]. Since *STK32b* and *MSX1* loci are in a close proximity in both the mouse and human genome, it is possible that the genetic manipulation of the *Msx1* locus caused aberrant expression of the *Stk32b* gene in mice. Nevertheless, we did not observe any interconnection between *Msx1* gene inactivation and *Stk32b* expression levels in the mouse intestine and thus the relationship between these two genes remains unclear. Furthermore, increased expression of the gene encoding transcription factor Sox17 was observed in both the $Msx1^{+}$ ectopic crypt cells and in the Apc-/Msx1-double-deficient epithelium (Figure 36). Previous studies described *Sox17* as a direct Wnt target gene[63] and a negative regulator of canonical Wnt signaling [335, 428]. It was also shown that ectopic expression of Sox transcription factors, including *Sox17*[106], converts differentiated cells into the stem cell-like state (reviewed in[306]). This is in accordance with our observation of decreased numbers of differentiated cells and increased expression of ISC markers in the Apc-/Msx1-double-deficient epithelium, where *Sox17* expression was elevated as well. It was previously suggested that Sox17 promotes degradation of β-catenin and Tcf4[335]. Since we observed increased *Sox17* gene expression in the Apc-/Msx1-double-deficient intestinal epithelium, we investigated β-catenin and Tcf4 expression using immunohistochemical staining. Our hypothesis was that loss of Msx1 could promote Sox17-mediated degradation of β-catenin and Tcf4 proteins. Nevertheless, we did not observe any changes in the β-catenin or Tcf4 protein production in the Msx1-deficient small intestinal epithelium or tumors. *SOX17* promoter hypermethylation and downregulation of the gene expression was found in human colorectal tumors[82, 333, 390, 427]. In contrast, we detected elevated levels of *SOX17* mRNA in human colorectal neoplasia as well as in the hyperproliferative mouse intestinal epithelium. However, although the *SOX17* promoter contains putative MSX1 binding sites, luciferase reporter assay did not reveal a direct regulation by MSX1.

The MSX1 function in human cancer cells, especially in CRC, has not been studied thoroughly and previous studies often brought contradictory suggestions of MSX1 function in cancer cells. Human CRC cells SW480 and SW620 with the *MSX1* gene disrupted seemed to grow faster than cells with the *MSX1* gene intact. Therefore we performed cell

cycle analysis and viability assay. Nevertheless, MSX1-deficient cells did not show any changes in proliferation or cell cycle progression. It was previously reported that Msx1 is required for the epithelial-mesenchymal transition (EMT) in the developing mouse heart[128]. EMT, a process when epithelial cells undergo biochemical and morphological changes to become migratory mesenchymal cells with invasive properties (reviewed in[156]), is activated in cancer cells during invasion and metastasis (reviewed in[101]). Therefore a "wound healing assay" was utilized to test the migratory properties of MSX1-deficient SW620 cells; however, *MSX1* inactivation had no effect on the ability of cells to migrate. Finally, to test the ability of MSX1-deficient cells to establish xenotransplants and grow *in vivo*, SW620 cells were subcutaneously injected into the lumbar back area of immunodeficient NSG™ mice. After 4 weeks, no differences were observed between dissected tumors derived from MSX1-deficient or control cells. We suppose that the lack of any (observable) differences between MSX1-deficient and MSX1-proficient SW620 cells may be attributed to the fact that SW620 cells are derived from a lymph node metastasis of a colorectal tumor. It is possible that these cells are already too "advanced", so *MSX1* inactivation does not cause any change in their proliferation capabilities. Because we observed the highest increase in *MSX1* expression in low grade dysplasia, we assume that MSX1 might have functions at early stages of tumor growth and its inactivation would not affect tumors in more advanced stages or even metastasis.

The gene expression profiling of human CRC cells SW620 revealed more than 200 genes with significantly altered expression upon *MSX1* gene disruption and some of these genes have been previously described as β-catenin target genes in SW480 cells[400]. As all of these genes were upregulated in MSX1-deficient SW620 cells, we supposed that MSX1 could function as a transcriptional repressor of genes activated by the Wnt/β-catenin signaling pathway. Nevertheless, we were not able to experimentally verify our hypothesis due to technical problems with chromatin sonication for the ChIP-seq analysis, which would probably be the best method for identification of functional MSX1-binding sites in the genome.

Finally, we tried to identify function of the tumor suppressor Hic1 in the mouse intestine using the mouse strain allowing *Hic1* inactivation throughout the intestinal epithelium. Inactivation of the *Hic1* gene in the mouse intestine led to a mild increase in numbers of Paneth, goblet, and enteroendocrine cells. This phenotype was probably caused by elevated expression of the Hic1 target gene *Atoh1*, which in the small intestine functions as the master regulator of secretory cell lineages[330], as it was previously described that

Atoh1 depletion results in loss of Paneth, goblet, and enteroendocrine cells[418]. We assume that the minor differences between the Hic1-deficient and the Hic1-proficient epithelium were caused by the utilization of the *Villin-Cre* mouse strain. The constitutively active Cre recombinase is expressed from the embryonic day 12.5 throughout the intestinal epithelium; however, in our previous experiments, we observed that some of the intestinal cells probably attenuate the Cre expression as the whole tissue appears to be a mosaic. We suppose that utilization of the strain expressing tamoxifen-regulated CreERT2 would provide more pronounced differences between the Hic1-deficient and Hic1-proficient tissue.

Since 2015, when we published our findings of Hic1 function in mouse intestines, only few articles dealing with Hic1 in intestinal tissue and tumors have been published. In 2016, Janeckova and co-workers described a study combining experiments performed on colorectal polyps and carcinomas from patients and bioinformatics analysis of publicly available datasets. Their analysis of *HIC1* gene expression and methylation indicated that *HIC1* is downregulated in premalignant stages of colorectal tumors due to methylation of its regulatory region. Of note, *HIC1* expression was specifically increased in a group of tumors sensitive to chemotherapy[147]. In consistence with these results, a year later Okazaki and colleagues discovered an important regulation of *HIC1* in patients with metastatic CRC, which could serve as a predictive mark for responsiveness to oxaliplatin-based chemotherapy. They found that the number of tandem repeat (TR) sequences, a 70 bp long sequence tandemly repeated in the *HIC1* promoter and associated with *HIC1* expression, indicates responsiveness to chemotherapy. HIC1 represses transcription of *SIRT1*, a histone deacetylase which indirectly activates the nucleotide excision repair (NER) pathway. In patients with less than four TRs in the *HIC1* promoter, the *HIC1* gene expression was not impaired, therefore *SIRT1* remained inhibited and could not activate the NER pathway, which contributed to oxaliplatin-sensitivity and better survival of treated patients[246]. Interestingly, Chen and colleagues reported that loss of *HIC1* methylation correlated with decreased migration of colorectal tumors[124]. Together, these data confirm that HIC1 is an important tumor suppressor, whose expression level can serve as a good marker determining the sensitivity of (colorectal) tumors to chemotherapy.

7 References

1 Aberle H, Bauer A, Stappert J, Kispert A, Kemler R. beta-catenin is a target for the ubiquitin-proteasome pathway. The EMBO journal 1997; 16: 3797-3804.

2 Aguilera O, Fraga MF, Ballestar E, Paz MF, Herranz M, Espada J *et al*. Epigenetic inactivation of the Wnt antagonist DICKKOPF-1 (DKK-1) gene in human colorectal cancer. Oncogene 2006; 25: 4116-4121.

3 Andreu P, Colnot S, Godard C, Gad S, Chafey P, Niwa-Kawakita M *et al*. Crypt-restricted proliferation and commitment to the Paneth cell lineage following Apc loss in the mouse intestine. Development 2005; 132: 1443-1451.

4 Andreu P, Peignon G, Slomianny C, Taketo MM, Colnot S, Robine S *et al*. A genetic study of the role of the Wnt/beta-catenin signalling in Paneth cell differentiation. Dev Biol 2008; 324: 288-296.

5 Aoki K, Taketo MM. Adenomatous polyposis coli (APC): a multi-functional tumor suppressor gene. J Cell Sci 2007; 120: 3327-3335.

6 Araki K, Ogata T, Kobayashi M, Yatani R. A morphological study on the histogenesis of human colorectal hyperplastic polyps. Gastroenterology 1995; 109: 1468-1474.

7 Ayabe T, Satchell DP, Wilson CL, Parks WC, Selsted ME, Ouellette AJ. Secretion of microbicidal alpha-defensins by intestinal Paneth cells in response to bacteria. Nat Immunol 2000; 1: 113-118.

8 Baeg GH, Lin X, Khare N, Baumgartner S, Perrimon N. Heparan sulfate proteoglycans are critical for the organization of the extracellular distribution of Wingless. Development 2001; 128: 87-94.

9 Bach A, Lallemand Y, Nicola MA, Ramos C, Mathis L, Maufras M *et al*. Msx1 is required for dorsal diencephalon patterning. Development 2003; 130: 4025-4036.

10 Bankaitis ED, Ha A, Kuo CJ, Magness ST. Reserve Stem Cells in Intestinal Homeostasis and Injury. Gastroenterology 2018; 155: 1348-1361.

11 Banziger C, Soldini D, Schutt C, Zipperlen P, Hausmann G, Basler K. Wntless, a conserved membrane protein dedicated to the secretion of Wnt proteins from signaling cells. Cell 2006; 125: 509-522.

12 Barker N, van Es JH, Kuipers J, Kujala P, van den Born M, Cozijnsen M *et al*. Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature 2007; 449: 1003-1007.

13 Barker N, Ridgway RA, van Es JH, van de Wetering M, Begthel H, van den Born M *et al*. Crypt stem cells as the cells-of-origin of intestinal cancer. Nature 2009; 457: 608-611.

14 Barker N, Clevers H. Leucine-rich repeat-containing G-protein-coupled receptors as markers of adult stem cells. Gastroenterology 2010; 138: 1681-1696.

15 Bass AJ, Lawrence MS, Brace LE, Ramos AH, Drier Y, Cibulskis K *et al*. Genomic sequencing of colorectal adenocarcinomas identifies a recurrent VTI1A-TCF7L2 fusion. Nat Genet 2011; 43: 964-968.

16 Batlle E, Henderson JT, Beghtel H, van den Born MM, Sancho E, Huls G *et al*. Beta-catenin and TCF mediate cell positioning in the intestinal epithelium by controlling the expression of EphB/ephrinB. Cell 2002; 111: 251-263.

17 Batts LE, Polk DB, Dubois RN, Kulessa H. Bmp signaling is required for intestinal growth and morphogenesis. Dev Dyn 2006; 235: 1563-1570.

18 Beckett K, Monier S, Palmer L, Alexandre C, Green H, Bonneil E *et al*. Drosophila S2 cells secrete wingless on exosome-like vesicles but the wingless gradient forms independently of exosomes. Traffic 2013; 14: 82-96.

19 Behrens J, von Kries JP, Kuhl M, Bruhn L, Wedlich D, Grosschedl R *et al*. Functional interaction of beta-catenin with the transcription factor LEF-1. Nature 1996; 382: 638-642.

20 Belenkaya TY, Wu Y, Tang X, Zhou B, Cheng L, Sharma YV *et al*. The retromer complex influences Wnt secretion by recycling wntless from endosomes to the trans-Golgi network. Dev Cell 2008; 14: 120-131.

21 Bendall AJ, Abate-Shen C. Roles for Msx and Dlx homeoproteins in vertebrate development. Gene 2000; 247: 17-31.

22 Bettington ML, Chetty R. Traditional serrated adenoma: an update. Hum Pathol 2015; 46: 933-938.

23 Bewick V, Cheek L, Ball J. Statistics review 12: survival analysis. Crit Care 2004; 8: 389-394.

24 Binato R, Alvarez Martinez CE, Pizzatti L, Robert B, Abdelhay E. SMAD 8 binding to mice Msx1 basal promoter is required for transcriptional activation. Biochem J 2006; 393: 141-150.

25 Bjerknes M, Cheng H. Clonal analysis of mouse intestinal epithelial progenitors. Gastroenterology 1999; 116: 7-14.

26 Bonds J, Pollan-White S, Xiang L, Mues G, D'Souza R. Is there a link between ovarian cancer and tooth agenesis? Eur J Med Genet 2014; 57: 235-239.

27 Bonito NA, Borley J, Wilhelm-Benartzi CS, Ghaem-Maghami S, Brown R. Epigenetic Regulation of the Homeobox Gene MSX1 Associates with Platinum-Resistant Disease in High-Grade Serous Epithelial Ovarian Cancer. Clin Cancer Res 2016; 22: 3097-3104.

28 Bossuyt W, Kazanjian A, De Geest N, Van Kelst S, De Hertogh G, Geboes K *et al*. Atonal homolog 1 is a tumor suppressor gene. PLoS Biol 2009; 7: e39.

29 Boulay G, Dubuissez M, Van Rechem C, Forget A, Helin K, Ayrault O *et al*. Hypermethylated in cancer 1 (HIC1) recruits polycomb repressive complex 2 (PRC2) to a subset of its target genes through interaction with human polycomb-like (hPCL) proteins. J Biol Chem 2012; 287: 10509-10524.

30 Brabletz T, Herrmann K, Jung A, Faller G, Kirchner T. Expression of nuclear beta-catenin and c-myc is correlated with tumor size but not with proliferative activity of colorectal adenomas. The American journal of pathology 2000; 156: 865-870.

31 Briggs KJ, Eberhart CG, Watkins DN. Just say no to ATOH: how HIC1 methylation might predispose medulloblastoma to lineage addiction. Cancer Res 2008; 68: 8654-8656.

32 Buechling T, Chaudhary V, Spirohn K, Weiss M, Boutros M. p24 proteins are required for secretion of Wnt ligands. EMBO Rep 2011; 12: 1265-1272.

33 Burda P, Padilla SM, Sarkar S, Emr SD. Retromer function in endosome-to-Golgi retrograde transport is regulated by the yeast Vps34 PtdIns 3-kinase. J Cell Sci 2002; 115: 3889-3900.

34 Burrows K, Antignano F, Bramhall M, Chenery A, Scheer S, Korinek V *et al*. The transcriptional repressor HIC1 regulates intestinal immune homeostasis. Mucosal Immunol 2017; 10: 1518-1528.

35 Cadigan KM, Peifer M. Wnt signaling from development to disease: insights from model systems. Cold Spring Harb Perspect Biol 2009; 1: a002881.

36 Cairnie AB, Millen BH. Fission of crypts in the small intestine of the irradiated mouse. Cell Tissue Kinet 1975; 8: 189-196.

37 Calvo R, Drabkin HA. Embryonic genes in cancer. Ann Oncol 2000; 11 Suppl 3: 207-218.

38 Carlton J, Bujny M, Peter BJ, Oorschot VM, Rutherford A, Mellor H *et al*. Sorting nexin-1 mediates tubular endosome-to-TGN transport through coincidence sensing of high-curvature membranes and 3-phosphoinositides. Curr Biol 2004; 14: 1791-1800.

39 Carter MG, Johns MA, Zeng X, Zhou L, Zink MC, Mankowski JL *et al*. Mice deficient in the candidate tumor suppressor gene Hic1 exhibit developmental defects of structures affected in the Miller-Dieker syndrome. Hum Mol Genet 2000; 9: 413-419.

40 Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. Bioinformatics 2010; 26: 2363-2367.

41 Catron KM, Zhang H, Marshall SC, Inostroza JA, Wilson JM, Abate C. Transcriptional repression by Msx-1 does not require homeodomain DNA-binding sites. Mol Cell Biol 1995; 15: 861-871.

42 Cavallo RA, Cox RT, Moline MM, Roose J, Polevoy GA, Clevers H *et al*. Drosophila Tcf and Groucho interact to repress Wingless signalling activity. Nature 1998; 395: 604-608.

43 Clarke RM. The effect of growth and of fasting on the number of villi and crypts in the small intestine of the albino rat. J Anat 1972; 112: 27-33.

44 Clevers H, Batlle E. EphB/EphrinB receptors and Wnt signaling in colorectal cancer. Cancer Res 2006; 66: 2-5.

45 Clevers H. The intestinal crypt, a prototype stem cell compartment. Cell 2013; 154: 274-284.

46 Cozier GE, Carlton J, McGregor AH, Gleeson PA, Teasdale RD, Mellor H *et al*. The phox homology (PX) domain-dependent, 3-phosphoinositide-mediated association of sorting nexin-1 with an early sorting endosomal compartment is required for its ability to regulate epidermal growth factor receptor degradation. J Biol Chem 2002; 277: 48730-48736.

47 Das S, Yu S, Sakamori R, Vedula P, Feng Q, Flores J *et al*. Rab8a vesicles regulate Wnt ligand delivery and Paneth cell maturation at the intestinal stem cell niche. Development 2015; 142: 2147-2162.

48 Davidson D. The function and evolution of Msx genes: pointers and paradoxes. Trends Genet 1995; 11: 405-411.

49 Davis H, Irshad S, Bansal M, Rafferty H, Boitsova T, Bardella C *et al*. Aberrant epithelial GREM1 expression initiates colonic tumorigenesis from cells outside the stem cell niche. Nat Med 2015; 21: 62-70.

50 De A. Wnt/Ca2+ signaling pathway: a brief overview. Acta Biochim Biophys Sin (Shanghai) 2011; 43: 745-756.

51 De Cecco L, Negri T, Brich S, Mauro V, Bozzi F, Dagrada G *et al*. Identification of a gene expression driven progression pathway in myxoid liposarcoma. Oncotarget 2014; 5: 5965-5977.

52 de Lau W, Barker N, Low TY, Koo BK, Li VS, Teunissen H *et al*. Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling. Nature 2011; 476: 293-297.

53 de Lau W, Kujala P, Schneeberger K, Middendorp S, Li VS, Barker N *et al*. Peyer's patch M cells derived from Lgr5(+) stem cells require SpiB and are induced by RankL in cultured "miniguts". Mol Cell Biol 2012; 32: 3639-3647.

54 De Robertis M, Massi E, Poeta ML, Carotti S, Morini S, Cecchetelli L *et al*. The AOM/DSS murine model for the study of colon carcinogenesis: From pathways to diagnosis and therapy studies. J Carcinog 2011; 10: 9.

55 Degirmenci B, Valenta T, Dimitrieva S, Hausmann G, Basler K. GLI1-expressing mesenchymal cells form the essential Wnt-secreting niche for colon stem cells. Nature 2018; 558: 449-453.

56 Dehennaut V, Leprince D. Implication of HIC1 (Hypermethylated In Cancer 1) in the DNA damage response. Bull Cancer 2009; 96: E66-72.

57 Deltour S, Pinte S, Guerardel C, Wasylyk B, Leprince D. The human candidate tumor suppressor gene HIC1 recruits CtBP through a degenerate GLDLSKK motif. Mol Cell Biol 2002; 22: 4890-4901.

58 Dhoot GK, Gustafsson MK, Ai X, Sun W, Standiford DM, Emerson CP, Jr. Regulation of Wnt signaling and embryo patterning by an extracellular sulfatase. Science 2001; 293: 1663-1666.

59 Dupasquier S, Abdel-Samad R, Glazer RI, Bastide P, Jay P, Joubert D *et al*. A new mechanism of SOX9 action to regulate PKCalpha expression in the intestine epithelium. J Cell Sci 2009; 122: 2191-2196.

60 Durand A, Donahue B, Peignon G, Letourneur F, Cagnard N, Slomianny C *et al*. Functional intestinal stem cells after Paneth cell ablation induced by the loss of transcription factor Math1 (Atoh1). Proc Natl Acad Sci U S A 2012; 109: 8965-8970.

61 Dwyer MA, Joseph JD, Wade HE, Eaton ML, Kunder RS, Kazmin D *et al*. WNT11 expression is induced by estrogen-related receptor alpha and beta-catenin and acts in an autocrine manner to increase cancer cell migration. Cancer Res 2010; 70: 9298-9308.

62 el Marjou F, Janssen KP, Chang BH, Li M, Hindie V, Chan L *et al*. Tissue-specific and inducible Cre-mediated recombination in the gut epithelium. Genesis 2004; 39: 186-193.

63 Engert S, Burtscher I, Liao WP, Dulev S, Schotta G, Lickert H. Wnt/beta-catenin signalling regulates Sox17 expression and is essential for organizer and endoderm formation in the mouse. Development 2013; 140: 3128-3138.

64 Fafilek B, Krausova M, Vojtechova M, Pospichalova V, Tumova L, Sloncova E *et al*. Troy, a tumor necrosis factor receptor family member, interacts with lgr5 to inhibit wnt signaling in intestinal stem cells. Gastroenterology 2013; 144: 381-391.

65 Farin HF, Van Es JH, Clevers H. Redundant Sources of Wnt Regulate Intestinal Stem Cells and Promote Formation of Paneth Cells. Gastroenterology 2012.

66 Farin HF, Jordens I, Mosa MH, Basak O, Korving J, Tauriello DV *et al*. Visualization of a short-range Wnt gradient in the intestinal stem-cell niche. Nature 2016; 530: 340-343.

67 Fedi P, Bafico A, Nieto Soria A, Burgess WH, Miki T, Bottaro DP *et al*. Isolation and biochemical characterization of the human Dkk-1 homologue, a novel inhibitor of mammalian Wnt signaling. J Biol Chem 1999; 274: 19465-19472.

68 Feil R, Wagner J, Metzger D, Chambon P. Regulation of Cre recombinase activity by mutated estrogen receptor ligand-binding domains. Biochem Biophys Res Commun 1997; 237: 752-757.

69 Feng XY, Wu XS, Wang JS, Zhang CM, Wang SL. Homeobox protein MSX-1 inhibits expression of bone morphogenetic protein 2, bone morphogenetic protein 4, and lymphoid enhancer-binding factor 1 via Wnt/beta-catenin signaling to prevent differentiation of dental mesenchymal cells during the late bell stage. Eur J Oral Sci 2018; 126: 1-12.

70 Feuerstein JD, Cheifetz AS. Crohn Disease: Epidemiology, Diagnosis, and Management. Mayo Clin Proc 2017; 92: 1088-1103.

71 Fleuriel C, Touka M, Boulay G, Guerardel C, Rood BR, Leprince D. HIC1 (Hypermethylated in Cancer 1) epigenetic silencing in tumors. Int J Biochem Cell Biol 2009; 41: 26-33.

72 Fodde R, Kuipers J, Rosenberg C, Smits R, Kielman M, Gaspar C *et al*. Mutations in the APC tumour suppressor gene cause chromosomal instability. Nat Cell Biol 2001; 3: 433-438.

73 Fodde R, Smits R. Disease model: familial adenomatous polyposis. Trends in molecular medicine 2001; 7: 369-373.

74 Fodde R, Smits R, Clevers H. APC, signal transduction and genetic instability in colorectal cancer. Nat Rev Cancer 2001; 1: 55-67.

75 Forbes SA, Bindal N, Beare D, Bamford S, Cole CG, Ward S *et al*. COSMIC: comprehensively exploring oncogenomics. Cancer Research 2016; 76.

76 Formeister EJ, Sionas AL, Lorance DK, Barkley CL, Lee GH, Magness ST. Distinct SOX9 levels differentially mark stem/progenitor populations and enteroendocrine cells of the small intestine epithelium. American journal of physiology Gastrointestinal and liver physiology 2009; 296: G1108-1118.

77 Forrester K, Almoguera C, Han K, Grizzle WE, Perucho M. Detection of high incidence of K-ras oncogenes during human colon tumorigenesis. Nature 1987; 327: 298-303.

78 Franch-Marro X, Wendler F, Guidato S, Griffith J, Baena-Lopez A, Itasaki N *et al*. Wingless secretion requires endosome-to-Golgi retrieval of Wntless/Evi/Sprinter by the retromer complex. Nat Cell Biol 2008; 10: 170-177.

79 Fre S, Hannezo E, Sale S, Huyghe M, Lafkas D, Kissel H *et al*. Notch lineages and activity in intestinal stem cells determined by a new set of knock-in mice. PLoS One 2011; 6: e25785.

80 Fujimitsu Y, Nakanishi H, Inada K, Yamachika T, Ichinose M, Fukami H *et al*. Development of aberrant crypt foci involves a fission mechanism as revealed by isolation of aberrant crypts. Jpn J Cancer Res 1996; 87: 1199-1203.

81 Furuta Y, Piston DW, Hogan BL. Bone morphogenetic proteins (BMPs) as regulators of dorsal forebrain development. Development 1997; 124: 2203-2212.

82 Galamb O, Kalmar A, Peterfia B, Csabai I, Bodor A, Ribli D *et al*. Aberrant DNA methylation of WNT pathway genes in the development and progression of CIMP-negative colorectal cancer. Epigenetics 2016; 11: 588-602.

83 Galiatsatos P, Foulkes WD. Familial adenomatous polyposis. The American journal of gastroenterology 2006; 101: 385-398.

84 Gasnereau I, Herr P, Chia PZ, Basler K, Gleeson PA. Identification of an endocytosis motif in an intracellular loop of Wntless protein, essential for its recycling and the control of Wnt protein signaling. J Biol Chem 2011; 286: 43324-43333.

85 Gaspar C, Fodde R. APC dosage effects in tumorigenesis and stem cell differentiation. The International journal of developmental biology 2004; 48: 377-386.

86 Gaspar C, Cardoso J, Franken P, Molenaar L, Morreau H, Moslein G *et al*. Cross-species comparison of human and mouse intestinal polyps reveals conserved mechanisms in adenomatous polyposis coli (APC)-driven tumorigenesis. The American journal of pathology 2008; 172: 1363-1380.

87 Gebhard A, Gebert A. Brush cells of the mouse intestine possess a specialized glycocalyx as revealed by quantitative lectin histochemistry. Further evidence for a sensory function. J Histochem Cytochem 1999; 47: 799-808.

88 Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S *et al*. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 2004; 5: R80.

89 Gerbe F, van Es JH, Makrini L, Brulin B, Mellitzer G, Robine S *et al*. Distinct ATOH1 and Neurog3 requirements define tuft cells as a new secretory cell type in the intestinal epithelium. J Cell Biol 2011; 192: 767-780.

90 Gonzalez R. World Health Organization (WHO) classification of colorectal carcinoma. PathologyOutlines.com website.
http://www.pathologyoutlines.com/topic/colontumorwhoclassification.html.
Accessed January 11th, 2019.

91 Gonzalez SM, Ferland LH, Robert B, Abdelhay E. Structural and functional analysis of mouse Msx1 gene promoter: sequence conservation with human MSX1 promoter points at potential regulatory elements. DNA Cell Biol 1998; 17: 561-572.

92 Goodman RM, Thombre S, Firtina Z, Gray D, Betts D, Roebuck J *et al*. Sprinter: a novel transmembrane protein required for Wg secretion and signaling. Development 2006; 133: 4901-4911.

93 Greco V, Hannus M, Eaton S. Argosomes: a potential vehicle for the spread of morphogens through epithelia. Cell 2001; 106: 633-645.

94 Gregorieff A, Pinto D, Begthel H, Destree O, Kielman M, Clevers H. Expression pattern of Wnt signaling components in the adult intestine. Gastroenterology 2005; 129: 626-638.

95 Grimm C, Sporle R, Schmid TE, Adler ID, Adamski J, Schughart K *et al*. Isolation and embryonic expression of the novel mouse gene Hic1, the homologue of HIC1, a candidate gene for the Miller-Dieker syndrome. Hum Mol Genet 1999; 8: 697-710.

96 Gross JC, Chaudhary V, Bartscherer K, Boutros M. Active Wnt proteins are secreted on exosomes. Nat Cell Biol 2012; 14: 1036-1045.

97 Guerardel C, Deltour S, Pinte S, Monte D, Begue A, Godwin AK *et al*. Identification in the human candidate tumor suppressor gene HIC-1 of a new major alternative TATA-less promoter positively regulated by p53. J Biol Chem 2001; 276: 3078-3089.

98 Gur G, Rubin C, Katz M, Amit I, Citri A, Nilsson J *et al*. LRIG1 restricts growth factor signaling by enhancing receptor ubiquitylation and degradation. The EMBO journal 2004; 23: 3270-3281.

99 Haft CR, de la Luz Sierra M, Bafford R, Lesniak MA, Barr VA, Taylor SI. Human orthologs of yeast vacuolar protein sorting proteins Vps26, 29, and 35: assembly into multimeric complexes. Mol Biol Cell 2000; 11: 4105-4116.

100 Han J, Ishii M, Bringas P, Jr., Maas RL, Maxson RE, Jr., Chai Y. Concerted action of Msx1 and Msx2 in regulating cranial neural crest cell differentiation during frontal bone development. Mech Dev 2007; 124: 729-745.

101 Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011; 144: 646-674.

102 Haramis AP, Begthel H, van den Born M, van Es J, Jonkheer S, Offerhaus GJ *et al*. De novo crypt formation and juvenile polyposis on BMP inhibition in mouse intestine. Science 2004; 303: 1684-1686.

103 Hardwick JC, Kodach LL, Offerhaus GJ, van den Brink GR. Bone morphogenetic protein signalling in colorectal cancer. Nat Rev Cancer 2008; 8: 806-812.

104 Hart MJ, de los Santos R, Albert IN, Rubinfeld B, Polakis P. Downregulation of beta-catenin by human Axin and its association with the APC tumor suppressor, beta-catenin and GSK3 beta. Curr Biol 1998; 8: 573-581.

105 Hattori N, Niwa T, Ishida T, Kobayashi K, Imai T, Mori A *et al*. Antibiotics suppress colon tumorigenesis through inhibition of aberrant DNA methylation in an azoxymethane and dextran sulfate sodium colitis model. Cancer science 2019; 110: 147-156.

106 He S, Kim I, Lim MS, Morrison SJ. Sox17 expression confers self-renewal potential and fetal stem cell characteristics upon adult hematopoietic progenitors. Genes Dev 2011; 25: 1613-1627.

107 He TC, Sparks AB, Rago C, Hermeking H, Zawel L, da Costa LT *et al*. Identification of c-MYC as a target of the APC pathway. Science 1998; 281: 1509-1512.

108 He XC, Zhang J, Tong WG, Tawfik O, Ross J, Scoville DH *et al*. BMP signaling inhibits intestinal stem cell self-renewal through suppression of Wnt-beta-catenin signaling. Nat Genet 2004; 36: 1117-1121.

109 Herr P, Hausmann G, Basler K. WNT secretion and signalling in human disease. Trends in molecular medicine 2012; 18: 483-493.

110 Hewitt JE, Clark LN, Ivens A, Williamson R. Structure and sequence of the human homeobox gene HOX7. Genomics 1991; 11: 670-678.

111 Hill RE, Jones PF, Rees AR, Sime CM, Justice MJ, Copeland NG *et al*. A new family of mouse homeo box-containing genes: molecular structure, chromosomal location, and developmental expression of Hox-7.1. Genes Dev 1989; 3: 26-37.

112 Holland PW. Evolution of homeobox genes. Wiley Interdiscip Rev Dev Biol 2013; 2: 31-45.

113 Holstein TW. The evolution of the Wnt pathway. Cold Spring Harb Perspect Biol 2012; 4: a007922.

114 Horazna M, Janeckova L, Svec J, Babosova O, Hrckulak D, Vojtechova M *et al*. Msx1 loss suppresses formation of the ectopic crypts developed in the Apc-deficient small intestinal epithelium. Sci Rep 2019; 9: 1629.

115 Horst D, Budczies J, Brabletz T, Kirchner T, Hlubek F. Invasion associated up-regulation of nuclear factor kappaB target genes in colorectal cancer. Cancer 2009; 115: 4946-4958.

116 Houzelstein D, Auda-Boucher G, Cheraud Y, Rouaud T, Blanc I, Tajbakhsh S *et al*. The homeobox gene Msx1 is expressed in a subset of somites, and in muscle progenitor cells migrating into the forelimb. Development 1999; 126: 2689-2701.

117 Hsieh JC, Kodjabachian L, Rebbert ML, Rattner A, Smallwood PM, Samos CH *et al*. A new secreted protein that binds to Wnt proteins and inhibits their activities. Nature 1999; 398: 431-436.

118 Huber MA, Kraut N, Beug H. Molecular requirements for epithelial-mesenchymal transition during tumor progression. Curr Opin Cell Biol 2005; 17: 548-558.

119 Humphries AC, Mlodzik M. From instruction to output: Wnt/PCP signaling in development and cancer. Curr Opin Cell Biol 2018; 51: 110-116.

120 Hung KE, Maricevich MA, Richard LG, Chen WY, Richardson MP, Kunin A *et al*. Development of a mouse model for sporadic and metastatic colon tumors and its use in assessing drug treatment. Proc Natl Acad Sci U S A 2010; 107: 1565-1570.

121 Chassaing B, Aitken JD, Malleshappa M, Vijay-Kumar M. Dextran sulfate sodium (DSS)-induced colitis in mice. Curr Protoc Immunol 2014; 104: Unit 15 25.

122 Chen E, Xu X, Liu T. Hereditary Nonpolyposis Colorectal Cancer and Cancer Syndromes: Recent Basic and Clinical Discoveries. J Oncol 2018; 2018: 3979135.

123 Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV *et al*. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics 2013; 14: 128.

124 Chen HC, Huang HY, Chen YL, Lee KD, Chu YR, Lin PY *et al*. Methylation of the Tumor Suppressor Genes HIC1 and RassF1A Clusters Independently From the Methylation of Polycomb Target Genes in Colon Cancer. Ann Surg Oncol 2017; 24: 578-585.

125 Chen W, Cooper TK, Zahnow CA, Overholtzer M, Zhao Z, Ladanyi M *et al*. Epigenetic and genetic loss of Hic1 function accentuates the role of p53 in tumorigenesis. Cancer Cell 2004; 6: 387-398.

126 Chen WY, Zeng X, Carter MG, Morrell CN, Chiu Yen RW, Esteller M *et al*. Heterozygous disruption of Hic1 predisposes mice to a gender-dependent spectrum of malignant tumors. Nat Genet 2003; 33: 197-202.

127 Chen YH, Ishii M, Sun J, Sucov HM, Maxson RE, Jr. Msx1 and Msx2 regulate survival of secondary heart field precursors and post-migratory proliferation of cardiac neural crest in the outflow tract. Dev Biol 2007; 308: 421-437.

128 Chen YH, Ishii M, Sucov HM, Maxson RE, Jr. Msx1 and Msx2 are required for endothelial-mesenchymal transformation of the atrioventricular cushions and patterning of the atrioventricular myocardium. BMC Dev Biol 2008; 8: 75.

129 Cheng H, Leblond CP. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. I. Columnar cell. The American journal of anatomy 1974; 141: 461-479.

130 Cheng H, Leblond CP. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. V. Unitarian Theory of the origin of the four epithelial cell types. The American journal of anatomy 1974; 141: 537-561.

131 Cheng H, Bjerknes M. Whole population cell kinetics and postnatal development of the mouse intestinal epithelium. Anat Rec 1985; 211: 420-426.

132 Cheng H, Bjerknes M, Amar J, Gardiner G. Crypt production in normal and diseased human colonic epithelium. Anat Rec 1986; 216: 44-48.

133 Cheng H, McCulloch C, Bjerknes M. Effects of 30% intestinal resection on whole population cell kinetics of mouse intestinal epithelium. Anat Rec 1986; 215: 35-41.

134 Chuang PT, McMahon AP. Vertebrate Hedgehog signalling modulated by induction of a Hedgehog-binding protein. Nature 1999; 397: 617-621.

135 Ikeda S, Kishida S, Yamamoto H, Murai H, Koyama S, Kikuchi A. Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3beta and beta-catenin and promotes GSK-3beta- dependent phosphorylation of beta-catenin. The EMBO journal 1998; 17: 1371-1384.

136 Ilyas M, Tomlinson IP, Rowan A, Pignatelli M, Bodmer WF. Beta-catenin mutations in cell lines established from human colorectal cancers. Proc Natl Acad Sci U S A 1997; 94: 10330-10334.

137 Ingersoll RG, Hetmanski J, Park JW, Fallin MD, McIntosh I, Wu-Chou YH *et al*. Association between genes on chromosome 4p16 and non-syndromic oral clefts in four populations. Eur J Hum Genet 2010; 18: 726-732.

138 Ireland H, Kemp R, Houghton C, Howard L, Clarke AR, Sansom OJ *et al*. Inducible Cre-mediated control of gene expression in the murine gastrointestinal tract: effect of loss of beta-catenin. Gastroenterology 2004; 126: 1236-1246.

139 Ireland H, Houghton C, Howard L, Winton DJ. Cellular inheritance of a Cre-activated reporter gene to determine Paneth cell longevity in the murine small intestine. Dev Dyn 2005; 233: 1332-1336.

140 Ishii M, Merrill AE, Chan YS, Gitelman I, Rice DP, Sucov HM *et al*. Msx2 and Twist cooperatively control the development of the neural crest-derived skeletogenic mesenchyme of the murine skull vault. Development 2003; 130: 6131-6142.

141 Ishii M, Han J, Yen HY, Sucov HM, Chai Y, Maxson RE, Jr. Combined deficiencies of Msx1 and Msx2 cause impaired patterning and survival of the cranial neural crest. Development 2005; 132: 4937-4950.

142 Itasaki N, Jones CM, Mercurio S, Rowe A, Domingos PM, Smith JC *et al*. Wise, a context-dependent activator and inhibitor of Wnt signalling. Development 2003; 130: 4295-4305.

143 Itzkovitz S, Lyubimova A, Blat IC, Maynard M, van Es J, Lees J *et al*. Single-molecule transcript counting of stem-cell markers in the mouse intestine. Nat Cell Biol 2012; 14: 106-114.

144 Ivens A, Flavin N, Williamson R, Dixon M, Bates G, Buckingham M *et al*. The human homeobox gene HOX7 maps to chromosome 4p16.1 and may be implicated in Wolf-Hirschhorn syndrome. Hum Genet 1990; 84: 473-476.

145 Jadhav U, Saxena M, O'Neill NK, Saadatpour A, Yuan GC, Herbert Z *et al*. Dynamic Reorganization of Chromatin Accessibility Signatures during Dedifferentiation of Secretory Precursors into Lgr5+ Intestinal Stem Cells. Cell Stem Cell 2017; 21: 65-77 e65.

146 Janeckova L, Pospichalova V, Fafilek B, Vojtechova M, Tureckova J, Dobes J *et al*. HIC1 Tumor Suppressor Loss Potentiates TLR2/NF-kappaB Signaling and Promotes Tissue Damage-Associated Tumorigenesis. Mol Cancer Res 2015; 13: 1139-1148.

147 Janeckova L, Kolar M, Svec J, Lanikova L, Pospichalova V, Baloghova N *et al*. HIC1 Expression Distinguishes Intestinal Carcinomas Sensitive to Chemotherapy. Transl Oncol 2016; 9: 99-107.

148 Janssen KP, Alberici P, Fsihi H, Gaspar C, Breukel C, Franken P *et al*. APC and oncogenic KRAS are synergistic in enhancing Wnt signaling in intestinal tumor formation and progression. Gastroenterology 2006; 131: 1096-1109.

149 Jen J, Powell SM, Papadopoulos N, Smith KJ, Hamilton SR, Vogelstein B *et al*. Molecular determinants of dysplasia in colorectal lesions. Cancer Res 1994; 54: 5523-5526.

150 Jenal M, Britschgi C, Fey MF, Tschan MP. Inactivation of the hypermethylated in cancer 1 tumour suppressor--not just a question of promoter hypermethylation? Swiss Med Wkly 2010; 140: w13106.

151 Jezewski PA, Vieira AR, Nishimura C, Ludwig B, Johnson M, O'Brien SE *et al*. Complete sequencing shows a role for MSX1 in non-syndromic cleft lip and palate. J Med Genet 2003; 40: 399-407.

152 Jho EH, Zhang T, Domon C, Joo CK, Freund JN, Costantini F. Wnt/beta-catenin/Tcf signaling induces the transcription of Axin2, a negative regulator of the signaling pathway. Mol Cell Biol 2002; 22: 1172-1183.

153 Jubb AM, Chalasani S, Frantz GD, Smits R, Grabsch HI, Kavi V *et al*. Achaete-scute like 2 (ascl2) is a target of Wnt signalling and is upregulated in intestinal neoplasia. Oncogene 2006; 25: 3445-3457.

154 Jumlongras D, Bei M, Stimson JM, Wang WF, DePalma SR, Seidman CE *et al*. A nonsense mutation in MSX1 causes Witkop syndrome. Am J Hum Genet 2001; 69: 67-74.

155 Kadowaki T, Wilder E, Klingensmith J, Zachary K, Perrimon N. The segment polarity gene porcupine encodes a putative multitransmembrane protein involved in Wingless processing. Genes Dev 1996; 10: 3116-3128.

156 Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. J Clin Invest 2009; 119: 1420-1428.

157 Karam SM. Lineage commitment and maturation of epithelial cells in the gut. Front Biosci 1999; 4: D286-298.

158 Kasparek P, Krausova M, Haneckova R, Kriz V, Zbodakova O, Korinek V *et al*. Efficient gene targeting of the Rosa26 locus in mouse zygotes using TALE nucleases. FEBS Lett 2014; 588: 3982-3988.

159 Kelly KF, Daniel JM. POZ for effect--POZ-ZF transcription factors in cancer and development. Trends Cell Biol 2006; 16: 578-587.

160 Kiecker C, Niehrs C. A morphogen gradient of Wnt/beta-catenin signalling regulates anteroposterior neural patterning in Xenopus. Development 2001; 128: 4189-4201.

161 Kim HJ, Rice DP, Kettunen PJ, Thesleff I. FGF-, BMP- and Shh-mediated signalling pathways in the regulation of cranial suture morphogenesis and calvarial bone development. Development 1998; 125: 1241-1251.

162 Kim S, Karin M. Role of TLR2-dependent inflammation in metastatic progression. Ann N Y Acad Sci 2011; 1217: 191-206.

163 Kim TH, Shivdasani RA. Genetic evidence that intestinal Notch functions vary regionally and operate through a common mechanism of Math1 repression. J Biol Chem 2011; 286: 11427-11433.

164 Kim TH, Escudero S, Shivdasani RA. Intact function of Lgr5 receptor-expressing intestinal stem cells in the absence of Paneth cells. Proc Natl Acad Sci U S A 2012; 109: 3932-3937.

165 Kinzler KW, Nilbert MC, Su LK, Vogelstein B, Bryan TM, Levy DB *et al*. Identification of FAP locus genes from chromosome 5q21. Science 1991; 253: 661-665.

166 Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. Cell 1996; 87: 159-170.

167 Kinzler KW, Vogelstein B. Cancer-susceptibility genes. Gatekeepers and caretakers. Nature 1997; 386: 761, 763.

168 Kitagawa M, Hatakeyama S, Shirane M, Matsumoto M, Ishida N, Hattori K *et al*. An F-box protein, FWD1, mediates ubiquitin-dependent proteolysis of beta-catenin. The EMBO journal 1999; 18: 2401-2410.

169 Kleinschmit A, Koyama T, Dejima K, Hayashi Y, Kamimura K, Nakato H. Drosophila heparan sulfate 6-O endosulfatase regulates Wingless morphogen gradient formation. Dev Biol 2010; 345: 204-214.

170 Koinuma K, Yamashita Y, Liu W, Hatanaka H, Kurashina K, Wada T *et al*. Epigenetic silencing of AXIN2 in colorectal carcinoma with microsatellite instability. Oncogene 2006; 25: 139-146.

171 Komekado H, Yamamoto H, Chiba T, Kikuchi A. Glycosylation and palmitoylation of Wnt-3a are coupled to produce an active form of Wnt-3a. Genes Cells 2007; 12: 521-534.

172 Koppens MA, Bounova G, Gargiulo G, Tanger E, Janssen H, Cornelissen-Steijger P *et al*. Deletion of Polycomb Repressive Complex 2 From Mouse Intestine Causes Loss of Stem Cells. Gastroenterology 2016; 151: 684-697 e612.

173 Korinek V, Barker N, Moerer P, van Donselaar E, Huls G, Peters PJ *et al*. Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. Nat Genet 1998; 19: 379-383.

174 Korkut C, Ataman B, Ramachandran P, Ashley J, Barria R, Gherbesi N *et al*. Trans-synaptic transmission of vesicular Wnt signals through Evi/Wntless. Cell 2009; 139: 393-404.

175 Kosinski C, Li VS, Chan AS, Zhang J, Ho C, Tsui WY *et al*. Gene expression patterns of human colon tops and basal crypts and BMP antagonists as intestinal stem cell niche factors. Proceedings of the National Academy of Sciences of the United States of America 2007; 104: 15418-15423.

176 Krausova M, Korinek V. Signal transduction pathways participating in homeostasis and malignant transformation of the intestinal tissue. Neoplasma 2012; 59: 708-718.

177 Krausova M, Korinek V. Wnt signaling in adult intestinal stem cells and cancer. Cellular signalling 2014; 26: 570-579.

178 Kriz V, Krausova M, Buresova P, Dobes J, Hrckulak D, Babosova O *et al*. Establishment of a tagged variant of Lgr4 receptor suitable for functional and expression studies in the mouse. Transgenic Res 2017; 26: 689-701.

179 Krupnik VE, Sharp JD, Jiang C, Robison K, Chickering TW, Amaravadi L *et al*. Functional and structural diversity of the human Dickkopf gene family. Gene 1999; 238: 301-313.

180 Kuhnert F, Davis CR, Wang HT, Chu P, Lee M, Yuan J *et al*. Essential requirement for Wnt signaling in proliferation of adult small intestine and colon revealed by adenoviral expression of Dickkopf-1. Proc Natl Acad Sci U S A 2004; 101: 266-271.

181 Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z *et al*. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016; 44: W90-97.

182 Kuraguchi M, Wang XP, Bronson RT, Rothenberg R, Ohene-Baah NY, Lund JJ *et al*. Adenomatous polyposis coli (APC) is required for normal development of skin and thymus. PLoS Genet 2006; 2: e146.

183 Lahn M, Paterson BM, Sundell K, Ma D. The role of protein kinase C-alpha (PKC-alpha) in malignancies of the gastrointestinal tract. Eur J Cancer 2004; 40: 10-20.

184 Lallemand Y, Nicola MA, Ramos C, Bach A, Clement CS, Robert B. Analysis of Msx1; Msx2 double mutants reveals multiple roles for Msx genes in limb development. Development 2005; 132: 3003-3014.

185 Langlands AJ, Almet AA, Appleton PL, Newton IP, Osborne JM, Nathke IS. Paneth Cell-Rich Regions Separated by a Cluster of Lgr5+ Cells Initiate Crypt Fission in the Intestinal Stem Cell Niche. PLoS Biol 2016; 14: e1002491.

186 Leibovitz A, Stinson JC, McCombs WB, 3rd, McCoy CE, Mazur KC, Mabry ND. Classification of human colorectal adenocarcinoma cell lines. Cancer Res 1976; 36: 4562-4569.

187 Leyns L, Bouwmeester T, Kim SH, Piccolo S, De Robertis EM. Frzb-1 is a secreted antagonist of Wnt signaling expressed in the Spemann organizer. Cell 1997; 88: 747-756.

188 Li X, Zhang Y, Kang H, Liu W, Liu P, Zhang J *et al*. Sclerostin binds to LRP5/6 and antagonizes canonical Wnt signaling. J Biol Chem 2005; 280: 19883-19887.

189 Liang J, Von den Hoff J, Lange J, Ren Y, Bian Z, Carels CE. MSX1 mutations and associated disease phenotypes: genotype-phenotype relations. Eur J Hum Genet 2016; 24: 1663-1670.

190 Lin K, Wang S, Julius MA, Kitajewski J, Moos M, Jr., Luyten FP. The cysteine-rich frizzled domain of Frzb-1 is required and sufficient for modulation of Wnt signaling. Proc Natl Acad Sci U S A 1997; 94: 11196-11200.

191 Lin X. Functions of heparan sulfate proteoglycans in cell signaling during development. Development 2004; 131: 6009-6021.

192 Liu C, Li Y, Semenov M, Han C, Baeg GH, Tan Y *et al*. Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. Cell 2002; 108: 837-847.

193 Liu W, Dong X, Mai M, Seelan RS, Taniguchi K, Krishnadath KK *et al*. Mutations in AXIN2 cause colorectal cancer with defective mismatch repair by activating beta-catenin/TCF signalling. Nat Genet 2000; 26: 146-147.

194 Liu W, Li H, Hong SH, Piszczek GP, Chen W, Rodgers GP. Olfactomedin 4 deletion induces colon adenocarcinoma in Apc(Min/+) mice. Oncogene 2016; 35: 5237-5247.

195 Logan CY, Nusse R. The Wnt signaling pathway in development and disease. Annu Rev Cell Dev Biol 2004; 20: 781-810.

196 Lopez-Garcia C, Klein AM, Simons BD, Winton DJ. Intestinal stem cell replacement follows a pattern of neutral drift. Science 2010; 330: 822-825.

197 Lukas J, Mazna P, Valenta T, Doubravska L, Pospichalova V, Vojtechova M *et al*. Dazap2 modulates transcription driven by the Wnt effector TCF-4. Nucleic Acids Res 2009; 37: 3007-3020.

198 Lustig B, Jerchow B, Sachs M, Weiler S, Pietsch T, Karsten U *et al*. Negative feedback loop of Wnt signaling through upregulation of conductin/axin2 in colorectal and liver tumors. Mol Cell Biol 2002; 22: 1184-1193.

199 Maas R, Chen YP, Bei M, Woo I, Satokata I. The role of Msx genes in mammalian development. Ann N Y Acad Sci 1996; 785: 171-181.

200 MacDonald BT, Tamai K, He X. Wnt/beta-catenin signaling: components, mechanisms, and diseases. Dev Cell 2009; 17: 9-26.

201 MacKenzie A, Purdie L, Davidson D, Collinson M, Hill RE. Two enhancer domains control early aspects of the complex expression pattern of Msx1. Mech Dev 1997; 62: 29-40.

202 Madison BB, Dunbar L, Qiao XT, Braunstein K, Braunstein E, Gumucio DL. Cis elements of the villin gene control expression in restricted domains of the vertical (crypt) and horizontal (duodenum, cecum) axes of the intestine. J Biol Chem 2002; 277: 33275-33283.

203 Madison BB, Braunstein K, Kuizon E, Portman K, Qiao XT, Gumucio DL. Epithelial hedgehog signals pattern the intestinal crypt-villus axis. Development 2005; 132: 279-289.

204 Malanchi I, Santamaria-Martinez A, Susanto E, Peng H, Lehr HA, Delaloye JF *et al*. Interactions between cancer stem cells and their niche govern metastatic colonization. Nature 2012; 481: 85-89.

205 Mao B, Wu W, Li Y, Hoppe D, Stannek P, Glinka A *et al*. LDL-receptor-related protein 6 is a receptor for Dickkopf proteins. Nature 2001; 411: 321-325.

206 Mao B, Wu W, Davidson G, Marhold J, Li M, Mechler BM *et al*. Kremen proteins are Dickkopf receptors that regulate Wnt/beta-catenin signalling. Nature 2002; 417: 664-667.

207 Mao B, Niehrs C. Kremen2 modulates Dickkopf2 activity during Wnt/LRP6 signaling. Gene 2003; 302: 179-183.

208 Marshman E, Booth C, Potten CS. The intestinal epithelial stem cell. Bioessays 2002; 24: 91-98.

209 Maskens AP, Dujardin-Loits RM. Kinetics of tissue proliferation in colorectal mucosa during post-natal growth. Cell Tissue Kinet 1981; 14: 467-477.

210 Massague J. TGFbeta signalling in context. Nat Rev Mol Cell Biol 2012; 13: 616-630.

211 Mattar MC, Lough D, Pishvaian MJ, Charabaty A. Current management of inflammatory bowel disease and colorectal cancer. Gastrointest Cancer Res 2011; 4: 53-61.

212 Maxson R, Ishii M, Merrill A. Msx genes in organogenesis and human disease. Advances in Developmental Biology and Biochemistry; Volume 13, 2003, Pages 43-68.

213 Mazieres J, He B, You L, Xu Z, Lee AY, Mikami I *et al*. Wnt inhibitory factor-1 is silenced by promoter hypermethylation in human lung cancer. Cancer Res 2004; 64: 4717-4720.

214 Medio M, Yeh E, Popelut A, Babajko S, Berdal A, Helms JA. Wnt/beta-catenin signaling and Msx1 promote outgrowth of the maxillary prominences. Frontiers in physiology 2012; 3: 375.

215 Melenovsky V, Benes J, Skaroupkova P, Sedmera D, Strnad H, Kolar M *et al*. Metabolic characterization of volume overload heart failure due to aorto-caval fistula in rats. Mol Cell Biochem 2011; 354: 83-96.

216 Menezes ME, Mitra A, Shevde LA, Samant RS. DNAJB6 governs a novel regulatory loop determining Wnt/beta-catenin signalling activity. Biochem J 2012; 444: 573-580.

217 Miller JR. The Wnts. Genome Biol 2002; 3: REVIEWS3001.

218 Miller KA, Barrow J, Collinson JM, Davidson S, Lear M, Hill RE *et al*. A highly conserved Wnt-dependent TCF4 binding site within the proximal enhancer of the anti-myogenic Msx1 gene supports expression within Pax3-expressing limb bud muscle precursor cells. Dev Biol 2007; 311: 665-678.

219 Minde DP, Anvarian Z, Rudiger SG, Maurice MM. Messing up disorder: how do missense mutations in the tumor suppressor protein APC lead to cancer? Mol Cancer 2011; 10: 101.

220 Miyaki M, Konishi M, Kikuchi-Yanoshita R, Enomoto M, Igari T, Tanaka K *et al*. Characteristics of somatic mutation of the adenomatous polyposis coli gene in colorectal tumors. Cancer Res 1994; 54: 3011-3020.

221 Miyoshi Y, Nagase H, Ando H, Horii A, Ichii S, Nakatsuru S *et al*. Somatic mutations of the APC gene in colorectal tumors: mutation cluster region in the APC gene. Hum Mol Genet 1992; 1: 229-233.

222 Mizokami Y, Egashira N, Takekoshi S, Itoh J, Itoh Y, Osamura RY *et al*. Expression of MSX1 in human normal pituitaries and pituitary adenomas. Endocrine pathology 2008; 19: 54-61.

223 Mohammad HP, Zhang W, Prevas HS, Leadem BR, Zhang M, Herman JG *et al*. Loss of a single Hic1 allele accelerates polyp formation in Apc(Delta716) mice. Oncogene 2011; 30: 2659-2669.

224 Molenaar M, van de Wetering M, Oosterwegel M, Peterson-Maduro J, Godsave S, Korinek V *et al*. XTcf-3 transcription factor mediates beta-catenin-induced axis formation in Xenopus embryos. Cell 1996; 86: 391-399.

225 Montgomery RK, Carlone DL, Richmond CA, Farilla L, Kranendonk ME, Henderson DE *et al*. Mouse telomerase reverse transcriptase (mTert) expression marks slowly cycling intestinal stem cells. Proc Natl Acad Sci U S A 2011; 108: 179-184.

226 Moore KA, Lemischka IR. Stem cells and their niches. Science 2006; 311: 1880-1885.

227 Mori-Akiyama Y, van den Born M, van Es JH, Hamilton SR, Adams HP, Zhang J *et al*. SOX9 is required for the differentiation of paneth cells in the intestinal epithelium. Gastroenterology 2007; 133: 539-546.

228 Morin PJ, Sparks AB, Korinek V, Barker N, Clevers H, Vogelstein B *et al*. Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC [see comments]. Science 1997; 275: 1787-1790.

229 Morroni M, Cangiotti AM, Cinti S. Brush cells in the human duodenojejunal junction: an ultrastructural study. J Anat 2007; 211: 125-131.

230 Moser AR, Pitot HC, Dove WF. A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. Science 1990; 247: 322-324.

231 Mulvaney J, Dabdoub A. Atoh1, an essential transcription factor in neurogenesis and intestinal and inner ear development: function, regulation, and context dependency. J Assoc Res Otolaryngol 2012; 13: 281-293.

232 Munoz J, Stange DE, Schepers AG, van de Wetering M, Koo BK, Itzkovitz S *et al*. The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent '+4' cell markers. The EMBO journal 2012; 31: 3079-3091.

233 Mustata RC, Van Loy T, Lefort A, Libert F, Strollo S, Vassart G *et al*. Lgr4 is required for Paneth cell differentiation and maintenance of intestinal stem cells ex vivo. EMBO Rep 2011; 12: 558-564.

234 Nagel S, Ehrentraut S, Meyer C, Kaufmann M, Drexler HG, MacLeod RA. Oncogenic deregulation of NKL homeobox gene MSX1 in mantle cell lymphoma. Leuk Lymphoma 2014; 55: 1893-1903.

235 Nagel S, Pommerenke C, Scherr M, Meyer C, Kaufmann M, Battmer K *et al*. NKL homeobox gene activities in hematopoietic stem cells, T-cell development and T-cell leukemia. PLoS One 2017; 12: e0171164.

236 Nagy A. Cre recombinase: the universal reagent for genome tailoring. Genesis 2000; 26: 99-109.

237 Neumann S, Coudreuse DY, van der Westhuyzen DR, Eckhardt ER, Korswagen HC, Schmitz G *et al*. Mammalian Wnt3a is released on lipoprotein particles. Traffic 2009; 10: 334-343.

238 Niehrs C. Head in the WNT: the molecular nature of Spemann's head organizer. Trends Genet 1999; 15: 314-319.

239 Nieminen P, Kotilainen J, Aalto Y, Knuutila S, Pirinen S, Thesleff I. MSX1 gene is deleted in Wolf-Hirschhorn syndrome patients with oligodontia. J Dent Res 2003; 82: 1013-1017.

240 Nishisho I, Nakamura Y, Miyoshi Y, Miki Y, Ando H, Horii A *et al*. Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. Science 1991; 253: 665-669.

241 Nordstrom U, Jessell TM, Edlund T. Progressive induction of caudal neural character by graded Wnt signaling. Nat Neurosci 2002; 5: 525-532.

242 Nusse R, Varmus HE. Wnt genes. Cell 1992; 69: 1073-1087.

243 Nusse R, Clevers H. Wnt/beta-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. Cell 2017; 169: 985-999.

244 Ogawa T, Kapadia H, Feng JQ, Raghow R, Peters H, D'Souza RN. Functional consequences of interactions between Pax9 and Msx1 genes in normal and abnormal tooth development. J Biol Chem 2006; 281: 18363-18369.

245 Ogi H, Suzuki K, Ogino Y, Kamimura M, Miyado M, Ying X *et al*. Ventral abdominal wall dysmorphogenesis of Msx1/Msx2 double-mutant mice. Anat Rec A Discov Mol Cell Evol Biol 2005; 284: 424-430.

246 Okazaki S, Schirripa M, Loupakis F, Cao S, Zhang W, Yang D *et al*. Tandem repeat variation near the HIC1 (hypermethylated in cancer 1) promoter predicts outcome of oxaliplatin-based chemotherapy in patients with metastatic colorectal cancer. Cancer 2017; 123: 4506-4514.

247 Ordonez-Moran P, Huelsken J. Lrig1: a new master regulator of epithelial stem cells. The EMBO journal 2012.

248 Orlando FA, Tan D, Baltodano JD, Khoury T, Gibbs JF, Hassid VJ *et al*. Aberrant crypt foci as precursors in colorectal cancer progression. J Surg Oncol 2008; 98: 207-213.

249 Oshima M, Oshima H, Kitagawa K, Kobayashi M, Itakura C, Taketo M. Loss of Apc heterozygosity and abnormal tissue building in nascent intestinal polyps in mice carrying a truncated Apc gene. Proc Natl Acad Sci U S A 1995; 92: 4482-4486.

250 Oster H, Leitges M. Protein kinase C alpha but not PKCzeta suppresses intestinal tumor formation in ApcMin/+ mice. Cancer Res 2006; 66: 6955-6963.

251 Pan CL, Baum PD, Gu M, Jorgensen EM, Clark SG, Garriga G. C. elegans AP-2 and retromer control Wnt signaling by regulating mig-14/Wntless. Dev Cell 2008; 14: 132-139.

252 Panakova D, Sprong H, Marois E, Thiele C, Eaton S. Lipoprotein particles are required for Hedgehog and Wingless signalling. Nature 2005; 435: 58-65.

253 Pani AM, Goldstein B. Direct visualization of a native Wnt in vivo reveals that a long-range Wnt gradient forms by extracellular dispersal. Elife 2018; 7.

254 Paoni NF, Feldman MW, Gutierrez LS, Ploplis VA, Castellino FJ. Transcriptional profiling of the transition from normal intestinal epithelia to adenomas and carcinomas in the APCMin/+ mouse. Physiol Genomics 2003; 15: 228-235.

255 Papadopoulos N, Lindblom A. Molecular basis of HNPCC: mutations of MMR genes. Human mutation 1997; 10: 89-99.

256 Parang B, Barrett CW, Williams CS. AOM/DSS Model of Colitis-Associated Cancer. Methods Mol Biol 2016; 1422: 297-307.

257 Park J, Park K, Kim S, Lee JH. Msx1 gene overexpression induces G1 phase cell arrest in human ovarian cancer cell line OVCAR3. Biochem Biophys Res Commun 2001; 281: 1234-1240.

258 Park K, Kim K, Rho SB, Choi K, Kim D, Oh SH *et al*. Homeobox Msx1 interacts with p53 tumor suppressor and inhibits tumor growth by inducing apoptosis. Cancer Res 2005; 65: 749-757.

259 Parris TZ, Aziz L, Kovacs A, Hajizadeh S, Nemes S, Semaan M *et al*. Clinical relevance of breast cancer-related genes as potential biomarkers for oral squamous cell carcinoma. BMC Cancer 2014; 14: 324.

260 Patterson AM, Watson AJM. Deciphering the Complex Signaling Systems That Regulate Intestinal Epithelial Cell Death Processes and Shedding. Front Immunol 2017; 8: 841.

261 Peeters T, Vantrappen G. The Paneth cell: a source of intestinal lysozyme. Gut 1975; 16: 553-558.

262 Peifer M, McCrea PD, Green KJ, Wieschaus E, Gumbiner BM. The vertebrate adhesive junction proteins beta-catenin and plakoglobin and the Drosophila segment polarity gene armadillo form a multigene family with similar properties. J Cell Biol 1992; 118: 681-691.

263 Peignon G, Durand A, Cacheux W, Ayrault O, Terris B, Laurent-Puig P *et al*. Complex interplay between beta-catenin signalling and Notch effectors in intestinal tumorigenesis. Gut 2011; 60: 166-176.

264 Pellegrinet L, Rodilla V, Liu Z, Chen S, Koch U, Espinosa L *et al*. Dll1- and dll4-mediated notch signaling are required for homeostasis of intestinal stem cells. Gastroenterology 2011; 140: 1230-1240 e1231-1237.

265 Perekatt AO, Shah PP, Cheung S, Jariwala N, Wu A, Gandhi V *et al*. SMAD4 Suppresses WNT-Driven Dedifferentiation and Oncogenesis in the Differentiated Gut Epithelium. Cancer Res 2018; 78: 4878-4890.

266 Phesse TJ, Durban VM, Sansom OJ. Defining key concepts of intestinal and epithelial cancer biology through the use of mouse models. Carcinogenesis 2017; 38: 953-965.

267 Piccolo S, Agius E, Leyns L, Bhattacharyya S, Grunz H, Bouwmeester T *et al*. The head inducer Cerberus is a multifunctional antagonist of Nodal, BMP and Wnt signals. Nature 1999; 397: 707-710.

268 Pinto D, Gregorieff A, Begthel H, Clevers H. Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. Genes Dev 2003; 17: 1709-1713.

269 Port F, Kuster M, Herr P, Furger E, Banziger C, Hausmann G *et al*. Wingless secretion promotes and requires retromer-dependent cycling of Wntless. Nat Cell Biol 2008; 10: 178-185.

270 Port F, Hausmann G, Basler K. A genome-wide RNA interference screen uncovers two p24 proteins as regulators of Wingless secretion. EMBO Rep 2011; 12: 1144-1152.

271 Pospichalova V, Tureckova J, Fafilek B, Vojtechova M, Krausova M, Lukas J *et al*. Generation of two modified mouse alleles of the Hic1 tumor suppressor gene. Genesis 2011; 49: 142-151.

272 Potten CS. Extreme sensitivity of some intestinal crypt cells to X and gamma irradiation. Nature 1977; 269: 518-521.

273 Potten CS, Hume WJ, Reid P, Cairns J. The segregation of DNA in epithelial stem cells. Cell 1978; 15: 899-906.

274 Potten CS, Booth C. The role of radiation-induced and spontaneous apoptosis in the homeostasis of the gastrointestinal epithelium: a brief review. Comparative biochemistry and physiology Part B, Biochemistry & molecular biology 1997; 118: 473-478.

275 Powell AE, Wang Y, Li Y, Poulin EJ, Means AL, Washington MK *et al*. The Pan-ErbB Negative Regulator Lrig1 Is an Intestinal Stem Cell Marker that Functions as a Tumor Suppressor. Cell 2012; 149: 146-158.

276 Powell DW, Mifflin RC, Valentich JD, Crowe SE, Saada JI, West AB. Myofibroblasts. II. Intestinal subepithelial myofibroblasts. The American journal of physiology 1999; 277: C183-201.

277 Powell SM, Zilz N, Beazer-Barclay Y, Bryan TM, Hamilton SR, Thibodeau SN *et al*. APC mutations occur early during colorectal tumorigenesis. Nature 1992; 359: 235-237.

278 Powell WC, Knox JD, Navre M, Grogan TM, Kittelson J, Nagle RB *et al*. Expression of the metalloproteinase matrilysin in DU-145 cells increases their invasive potential in severe combined immunodeficient mice. Cancer Res 1993; 53: 417-422.

279 Preston SL, Wong WM, Chan AO, Poulsom R, Jeffery R, Goodlad RA *et al*. Bottom-up histogenesis of colorectal adenomas: origin in the monocryptal adenoma and initial expansion by crypt fission. Cancer Res 2003; 63: 3819-3825.

280 Qi Z, Li Y, Zhao B, Xu C, Liu Y, Li H *et al*. BMP restricts stemness of intestinal Lgr5(+) stem cells by directly suppressing their signature genes. Nat Commun 2017; 8: 13824.

281 Qiu W, Hu Y, Andersen TE, Jafari A, Li N, Chen W *et al*. Tumor necrosis factor receptor superfamily member 19 (TNFRSF19) regulates differentiation fate of human mesenchymal (stromal) stem cells through canonical Wnt signaling and C/EBP. J Biol Chem 2010; 285: 14438-14449.

282 Quyn AJ, Appleton PL, Carey FA, Steele RJ, Barker N, Clevers H *et al*. Spindle orientation bias in gut epithelial stem cell compartments is lost in precancerous tissue. Cell Stem Cell 2010; 6: 175-181.

283 Revet I, Huizenga G, Koster J, Volckmann R, van Sluis P, Versteeg R *et al*. MSX1 induces the Wnt pathway antagonist genes DKK1, DKK2, DKK3, and SFRP1 in neuroblastoma cells, but does not block Wnt3 and Wnt5A signalling to DVL3. Cancer Lett 2010; 289: 195-207.

284 Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. Nature 2001; 414: 105-111.

285 Rhodes JM, Campbell BJ. Inflammation and colorectal cancer: IBD-associated and sporadic cancer compared. Trends in molecular medicine 2002; 8: 10-16.

286 Riccio O, van Gijn ME, Bezdek AC, Pellegrinet L, van Es JH, Zimber-Strobl U *et al*. Loss of intestinal crypt progenitor cells owing to inactivation of both Notch1 and Notch2 is accompanied by derepression of CDK inhibitors p27Kip1 and p57Kip2. EMBO Rep 2008; 9: 377-383.

287 Ritsma L, Ellenbroek SIJ, Zomer A, Snippert HJ, de Sauvage FJ, Simons BD *et al*. Intestinal crypt homeostasis revealed at single-stem-cell level by in vivo live imaging. Nature 2014; 507: 362-365.

288 Robert B, Sassoon D, Jacq B, Gehring W, Buckingham M. Hox-7, a mouse homeobox gene with a novel pattern of expression during embryogenesis. The EMBO journal 1989; 8: 91-100.

289 Roberts DJ, Johnson RL, Burke AC, Nelson CE, Morgan BA, Tabin C. Sonic hedgehog is an endodermal signal inducing Bmp-4 and Hox genes during induction and regionalization of the chick hindgut. Development 1995; 121: 3163-3174.

290 Rodrigues NR, Rowan A, Smith ME, Kerr IB, Bodmer WF, Gannon JV *et al*. p53 mutations in colorectal cancer. Proc Natl Acad Sci U S A 1990; 87: 7555-7559.

291 Roche KC, Gracz AD, Liu XF, Newton V, Akiyama H, Magness ST. SOX9 maintains reserve stem cells and preserves radioresistance in mouse small intestine. Gastroenterology 2015; 149: 1553-1563 e1510.

292 Rood BR, Leprince D. Deciphering HIC1 control pathways to reveal new avenues in cancer therapeutics. Expert Opin Ther Targets 2013; 17: 811-827.

293 Roose J, Molenaar M, Peterson J, Hurenkamp J, Brantjes H, Moerer P *et al*. The Xenopus Wnt effector XTcf-3 interacts with Groucho-related transcriptional repressors. Nature 1998; 395: 608-612.

294 Roth S, Franken P, Sacchetti A, Kremer A, Anderson K, Sansom O *et al*. Paneth cells in intestinal homeostasis and tissue injury. PLoS One 2012; 7: e38965.

295 Rothenberg ME, Nusse Y, Kalisky T, Lee JJ, Dalerba P, Scheeren F *et al*. Identification of a cKit(+) Colonic Crypt Base Secretory Cell That Supports Lgr5(+) Stem Cells in Mice. Gastroenterology 2012.

296 Rothenberg ME, Nusse Y, Kalisky T, Lee JJ, Dalerba P, Scheeren F *et al*. Identification of a cKit(+) colonic crypt base secretory cell that supports Lgr5(+) stem cells in mice. Gastroenterology 2012; 142: 1195-1205 e1196.

297 Rubinfeld B, Albert I, Porfiri E, Fiol C, Munemitsu S, Polakis P. Binding of GSK3beta to the APC-beta-catenin complex and regulation of complex assembly. Science 1996; 272: 1023-1026.

298 Saadi I, Das P, Zhao M, Raj L, Ruspita I, Xia Y *et al*. Msx1 and Tbx2 antagonistically regulate Bmp4 expression during the bud-to-cap stage transition in tooth development. Development 2013; 140: 2697-2702.

299 Sailer MH, Gerber A, Tostado C, Hutter G, Cordier D, Mariani L *et al*. Non-invasive neural stem cells become invasive in vitro by combined FGF2 and BMP4 signaling. J Cell Sci 2013; 126: 3533-3540.

300 Samuels Y, Wang Z, Bardelli A, Silliman N, Ptak J, Szabo S *et al*. High frequency of mutations of the PIK3CA gene in human cancers. Science 2004; 304: 554.

301 San Roman AK, Jayewickreme CD, Murtaugh LC, Shivdasani RA. Wnt secretion from epithelial cells and subepithelial myofibroblasts is not required in the mouse intestinal stem cell niche in vivo. Stem Cell Reports 2014; 2: 127-134.

302 Sangiorgi E, Capecchi MR. Bmi1 is expressed in vivo in intestinal stem cells. Nat Genet 2008; 40: 915-920.

303 Sancho A, Vandersmissen I, Craps S, Luttun A, Groll J. A new strategy to measure intercellular adhesion forces in mature cell-cell contacts. Sci Rep 2017; 7: 46152.

304 Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods 2014; 11: 783-784.

305 Sansom OJ, Reed KR, Hayes AJ, Ireland H, Brinkmann H, Newton IP *et al*. Loss of Apc in vivo immediately perturbs Wnt signaling, differentiation, and migration. Genes Dev 2004; 18: 1385-1390.

306 Sarkar A, Hochedlinger K. The sox family of transcription factors: versatile regulators of stem and progenitor cell fate. Cell Stem Cell 2013; 12: 15-30.

307 Sarrazin S, Lamanna WC, Esko JD. Heparan sulfate proteoglycans. Cold Spring Harb Perspect Biol 2011; 3.

308 Sasaki N, Sachs N, Wiebrands K, Ellenbroek SI, Fumagalli A, Lyubimova A *et al*. Reg4+ deep crypt secretory cells function as epithelial niche for Lgr5+ stem cells in colon. Proc Natl Acad Sci U S A 2016; 113: E5399-5407.

309 Sato T, Vries RG, Snippert HJ, van de Wetering M, Barker N, Stange DE *et al*. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature 2009; 459: 262-265.

310 Sato T, van Es JH, Snippert HJ, Stange DE, Vries RG, van den Born M *et al*. Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. Nature 2011; 469: 415-418.

311 Satokata I, Maas R. Msx1 deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development. Nat Genet 1994; 6: 348-356.

312 Satokata I, Ma L, Ohshima H, Bei M, Woo I, Nishizawa K *et al*. Msx2 deficiency in mice causes pleiotropic defects in bone growth and ectodermal organ formation. Nat Genet 2000; 24: 391-395.

313 Sauer B, Henderson N. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. Proc Natl Acad Sci U S A 1988; 85: 5166-5170.

314 Seaman MN, McCaffery JM, Emr SD. A membrane coat complex essential for endosome-to-Golgi retrograde transport in yeast. J Cell Biol 1998; 142: 665-681.

315 Seaman MN. Identification of a novel conserved sorting motif required for retromer-mediated endosome-to-TGN retrieval. J Cell Sci 2007; 120: 2378-2389.

316 Segditsas S, Tomlinson I. Colorectal cancer and genetic alterations in the Wnt pathway. Oncogene 2006; 25: 7531-7537.

317 Semenov MV, Tamai K, Brott BK, Kuhl M, Sokol S, He X. Head inducer Dickkopf-1 is a ligand for Wnt coreceptor LRP6. Curr Biol 2001; 11: 951-961.

318 Serralbo O, Marcelle C. Migrating cells mediate long-range WNT signaling. Development 2014; 141: 2057-2063.

319 Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T *et al*. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 2014; 343: 84-87.

320 Shames DS, Girard L, Gao B, Sato M, Lewis CM, Shivapurkar N *et al*. A genome-wide screen for promoter methylation in lung cancer identifies novel methylation markers for multiple malignancies. PLoS medicine 2006; 3: e486.

321 Shenoy AK, Fisher RC, Butterworth EA, Pi L, Chang LJ, Appelman HD *et al*. Transition from colitis to cancer: high Wnt activity sustains the tumor-initiating potential of colon cancer stem cell precursors. Cancer Res 2012; 72: 5091-5100.

322 Shim C, Zhang W, Rhee CH, Lee JH. Profiling of differentially expressed genes in human primary cervical cancer by complementary DNA expression array. Clin Cancer Res 1998; 4: 3045-3050.

323 Shimeld SM, McKay IJ, Sharpe PT. The murine homeobox gene Msx-3 shows highly restricted expression in the developing neural tube. Mech Dev 1996; 55: 201-210.

324 Shimizu Y, Ikeda S, Fujimori M, Kodama S, Nakahara M, Okajima M *et al*. Frequent alterations in the Wnt signaling pathway in colorectal cancer with microsatellite instability. Genes Chromosomes Cancer 2002; 33: 73-81.

325 Shtutman M, Zhurinsky J, Simcha I, Albanese C, D'Amico M, Pestell R *et al*. The cyclin D1 gene is a target of the beta-catenin/LEF-1 pathway. Proc Natl Acad Sci U S A 1999; 96: 5522-5527.

326 Schmitt M, Metzger M, Gradl D, Davidson G, Orian-Rousseau V. CD44 functions in Wnt signaling by regulating LRP6 localization and activation. Cell Death Differ 2015; 22: 677-689.

327 Schuijers J, Junker JP, Mokry M, Hatzis P, Koo BK, Sasselli V *et al*. Ascl2 acts as an R-spondin/Wnt-responsive switch to control stemness in intestinal crypts. Cell Stem Cell 2015; 16: 158-170.

328 Schwitalla S, Fingerle AA, Cammareri P, Nebelsiek T, Goktuna SI, Ziegler PK *et al*. Intestinal Tumorigenesis Initiated by Dedifferentiation and Acquisition of Stem-Cell-like Properties. Cell 2012.

329 Schwitalla S, Fingerle AA, Cammareri P, Nebelsiek T, Goktuna SI, Ziegler PK *et al*. Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. Cell 2013; 152: 25-38.

330 Siebel C, Lendahl U. Notch Signaling in Development, Tissue Homeostasis, and Disease. Physiol Rev 2017; 97: 1235-1294.

331 Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. CA: a cancer journal for clinicians 2012; 62: 10-29.

332 Silhankova M, Port F, Harterink M, Basler K, Korswagen HC. Wnt signalling requires MTM-6 and MTM-9 myotubularin lipid-phosphatase function in Wnt-producing cells. The EMBO journal 2010; 29: 4094-4105.

333 Silva AL, Dawson SN, Arends MJ, Guttula K, Hall N, Cameron EA *et al*. Boosting Wnt activity during colorectal cancer progression through selective hypermethylation of Wnt signaling antagonists. BMC Cancer 2014; 14: 891.

334 Singh R, Zorron Cheng Tao Pu L, Koay D, Burt A. Sessile serrated adenoma/polyps: Where are we at in 2016? World J Gastroenterol 2016; 22: 7754-7759.

335 Sinner D, Kordich JJ, Spence JR, Opoka R, Rankin S, Lin SC *et al*. Sox17 and Sox4 differentially regulate beta-catenin/T-cell factor activity and proliferation of colon carcinoma cells. Mol Cell Biol 2007; 27: 7802-7815.

336 Slattery ML, Herrick JS, Bondurant KL, Wolff RK. Toll-like receptor genes and their association with colon and rectal cancer development and prognosis. Int J Cancer 2012; 130: 2974-2980.

337 Smith AJ, Stern HS, Penner M, Hay K, Mitri A, Bapat BV *et al*. Somatic APC and K-ras codon 12 mutations in aberrant crypt foci from human colons. Cancer Res 1994; 54: 5527-5530.

338 Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 2004; 3: Article3.

339 Snippert HJ, van der Flier LG, Sato T, van Es JH, van den Born M, Kroon-Veenboer C *et al*. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. Cell 2010; 143: 134-144.

340 Snippert HJ, Schepers AG, van Es JH, Simons BD, Clevers H. Biased competition between Lgr5 intestinal stem cells driven by oncogenic mutation induces clonal expansion. EMBO Rep 2014; 15: 62-69.

341 Snover DC. Update on the serrated pathway to colorectal carcinoma. Hum Pathol 2011; 42: 1-10.

342 Solanas G, Cortina C, Sevillano M, Batlle E. Cleavage of E-cadherin by ADAM10 mediates epithelial cell sorting downstream of EphB signalling. Nat Cell Biol 2011; 13: 1100-1107.

343 Song K, Wang Y, Sassoon D. Expression of Hox-7.1 in myoblasts inhibits terminal differentiation and induces cell transformation. Nature 1992; 360: 477-481.

344 Srinivasan T, Than EB, Bu P, Tung KL, Chen KY, Augenlicht L *et al*. Notch signalling regulates asymmetric division and inter-conversion between lgr5 and bmi1 expressing intestinal stem cells. Sci Rep 2016; 6: 26069.

345 St Clair WH, Osborne JW. Crypt fission and crypt number in the small and large bowel of postnatal rats. Cell Tissue Kinet 1985; 18: 255-262.

346 Stanganello E, Scholpp S. Role of cytonemes in Wnt transport. J Cell Sci 2016; 129: 665-672.

347 Sternini C, Anselmi L, Rozengurt E. Enteroendocrine cells: a site of 'taste' in gastrointestinal chemosensing. Curr Opin Endocrinol Diabetes Obes 2008; 15: 73-78.

348 Storey JD, Tibshirani R. Statistical methods for identifying differentially expressed genes in DNA microarrays. Methods Mol Biol 2003; 224: 149-157.

349 Su LK, Kinzler KW, Vogelstein B, Preisinger AC, Moser AR, Luongo C *et al*. Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene. Science 1992; 256: 668-670.

350 Sun M, Song H, Wang S, Zhang C, Zheng L, Chen F *et al*. Integrated analysis identifies microRNA-195 as a suppressor of Hippo-YAP pathway in colorectal cancer. J Hematol Oncol 2017; 10: 79.

351 Suzuki A, Sekiya S, Gunshima E, Fujii S, Taniguchi H. EGF signaling activates proliferation and blocks apoptosis of mouse and human intestinal stem/progenitor cells in long-term monolayer cell culture. Lab Invest 2010; 90: 1425-1436.

352 Suzuki H, Gabrielson E, Chen W, Anbazhagan R, van Engeland M, Weijenberg MP *et al*. A genomic screen for genes upregulated by demethylation and histone deacetylase inhibition in human colorectal cancer. Nat Genet 2002; 31: 141-149.

353 Taciak B, Pruszynska I, Kiraga L, Bialasek M, Krol M. Wnt signaling pathway in development and cancer. J Physiol Pharmacol 2018; 69.

354 Taipale J, Beachy PA. The Hedgehog and Wnt signalling pathways in cancer. Nature 2001; 411: 349-354.

355 Takada R, Satomi Y, Kurata T, Ueno N, Norioka S, Kondoh H *et al*. Monounsaturated fatty acid modification of Wnt protein: its role in Wnt secretion. Dev Cell 2006; 11: 791-801.

356 Takahashi M, Nakamura Y, Obama K, Furukawa Y. Identification of SP5 as a downstream gene of the beta-catenin/Tcf pathway and its enhanced expression in human colon cancer. Int J Oncol 2005; 27: 1483-1487.

357 Takahashi T, Guron C, Shetty S, Matsui H, Raghow R. A minimal murine Msx-1 gene promoter. Organization of its cis-regulatory motifs and their role in transcriptional activation in cells in culture and in transgenic mice. J Biol Chem 1997; 272: 22667-22678.

358 Takeda N, Jain R, LeBoeuf MR, Wang Q, Lu MM, Epstein JA. Interconversion between intestinal stem cell populations in distinct niches. Science 2011; 334: 1420-1424.

359 Taketo MM, Edelmann W. Mouse models of colon cancer. Gastroenterology 2009; 136: 780-798.

360 Tan BT, Park CY, Ailles LE, Weissman IL. The cancer stem cell hypothesis: a work in progress. Lab Invest 2006; 86: 1203-1207.

361 Tanaka T, Kohno H, Suzuki R, Yamada Y, Sugie S, Mori H. A novel inflammation-related mouse colon carcinogenesis model induced by azoxymethane and dextran sodium sulfate. Cancer science 2003; 94: 965-973.

362 Tang W, Dodge M, Gundapaneni D, Michnoff C, Roth M, Lum L. A genome-wide RNAi screen for Wnt/beta-catenin pathway components identifies unexpected roles for TCF transcription factors in cancer. Proceedings of the National Academy of Sciences of the United States of America 2008; 105: 9697-9702.

363 Tao H, Guo L, Chen L, Qiao G, Meng X, Xu B *et al*. MSX1 inhibits cell migration and invasion through regulating the Wnt/beta-catenin pathway in glioblastoma. Tumour Biol 2016; 37: 1097-1104.

364 Temtamy SA, Aglan MS, Valencia M, Cocchi G, Pacheco M, Ashour AM *et al*. Long interspersed nuclear element-1 (LINE1)-mediated deletion of EVC, EVC2, C4orf6, and STK32B in Ellis-van Creveld syndrome with borderline intelligence. Human mutation 2008; 29: 931-938.

365 Tetsu O, McCormick F. Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. Nature 1999; 398: 422-426.

366 Tetsu O, McCormick F. Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. Nature 1999; 398: 422-426.

367 Tetteh PW, Basak O, Farin HF, Wiebrands K, Kretzschmar K, Begthel H *et al*. Replacement of Lost Lgr5-Positive Stem Cells through Plasticity of Their Enterocyte-Lineage Daughters. Cell Stem Cell 2016; 18: 203-213.

368 Thomsen ER, Mich JK, Yao Z, Hodge RD, Doyle AM, Jang S *et al*. Fixed single-cell transcriptomic characterization of human radial glial diversity. Nat Methods 2016; 13: 87-93.

369 Tian H, Biehs B, Warming S, Leong KG, Rangell L, Klein OD *et al*. A reserve stem cell population in small intestine renders Lgr5-positive cells dispensable. Nature 2011; 478: 255-259.

370 Torlakovic EE, Gomez JD, Driman DK, Parfitt JR, Wang C, Benerjee T *et al*. Sessile serrated adenoma (SSA) vs. traditional serrated adenoma (TSA). Am J Surg Pathol 2008; 32: 21-29.

371 Totafurno J, Bjerknes M, Cheng H. The crypt cycle. Crypt and villus production in the adult intestinal epithelium. Biophys J 1987; 52: 279-294.

372 Tumova L, Pombinho AR, Vojtechova M, Stancikova J, Gradl D, Krausova M *et al*. Monensin inhibits canonical Wnt signaling in human colorectal cancer cells and suppresses tumor growth in multiple intestinal neoplasia mice. Molecular cancer therapeutics 2014; 13: 812-822.

373 Ungaro R, Mehandru S, Allen PB, Peyrin-Biroulet L, Colombel JF. Ulcerative colitis. Lancet 2017; 389: 1756-1770.

374 Valenta T, Lukas J, Doubravska L, Fafilek B, Korinek V. HIC1 attenuates Wnt signaling by recruitment of TCF-4 and beta-catenin to the nuclear bodies. The EMBO journal 2006; 25: 2326-2337.

375 van Amerongen R, Nusse R. Towards an integrated view of Wnt signaling in development. Development 2009; 136: 3205-3214.

376 van de Wetering M, Sancho E, Verweij C, de Lau W, Oving I, Hurlstone A *et al*. The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. Cell 2002; 111: 241-250.

377 van der Flier LG, van Gijn ME, Hatzis P, Kujala P, Haegebarth A, Stange DE *et al*. Transcription factor achaete scute-like 2 controls intestinal stem cell fate. Cell 2009; 136: 903-912.

378 Van Der Kraak L, Gros P, Beauchemin N. Colitis-associated colon cancer: Is it in your genes? World J Gastroenterol 2015; 21: 11688-11699.

379 van Dop WA, Uhmann A, Wijgerde M, Sleddens-Linkels E, Heijmans J, Offerhaus GJ *et al*. Depletion of the colonic epithelial precursor cell compartment upon conditional activation of the hedgehog pathway. Gastroenterology 2009; 136: 2195-2203 e2191-2197.

380 van Es JH, van Gijn ME, Riccio O, van den Born M, Vooijs M, Begthel H *et al*. Notch/gamma-secretase inhibition turns proliferative cells in intestinal crypts and adenomas into goblet cells. Nature 2005; 435: 959-963.

381 van Es JH, Sato T, van de Wetering M, Lyubimova A, Nee AN, Gregorieff A *et al*. Dll1+ secretory progenitor cells revert to stem cells upon crypt damage. Nat Cell Biol 2012; 14: 1099-1104.

382 Van Landeghem L, Santoro MA, Krebs AE, Mah AT, Dehmer JJ, Gracz AD *et al*. Activation of two distinct Sox9-EGFP-expressing intestinal stem cell populations during crypt regeneration after irradiation. American journal of physiology Gastrointestinal and liver physiology 2012; 302: G1111-1132.

383 Van Landeghem L, Santoro MA, Mah AT, Krebs AE, Dehmer JJ, McNaughton KK *et al*. IGF1 stimulates crypt expansion via differential activation of 2 intestinal stem cell populations. FASEB J 2015; 29: 2828-2842.

384 Van Rechem C, Rood BR, Touka M, Pinte S, Jenal M, Guerardel C *et al*. Scavenger chemokine (CXC motif) receptor 7 (CXCR7) is a direct target gene of HIC1 (hypermethylated in cancer 1). J Biol Chem 2009; 284: 20927-20935.

385 Van Rechem C, Boulay G, Pinte S, Stankovic-Valentin N, Guerardel C, Leprince D. Differential regulation of HIC1 target genes by CtBP and NuRD, via an acetylation/SUMOylation switch, in quiescent versus proliferating cells. Mol Cell Biol 2010; 30: 4045-4059.

386 Vastardis H, Karimbux N, Guthua SW, Seidman JG, Seidman CE. A human MSX1 homeodomain missense mutation causes selective tooth agenesis. Nat Genet 1996; 13: 417-421.

387 Veeman MT, Axelrod JD, Moon RT. A second canon. Functions and mechanisms of beta-catenin-independent Wnt signaling. Dev Cell 2003; 5: 367-377.

388 Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M *et al*. Genetic alterations during colorectal-tumor development. N Engl J Med 1988; 319: 525-532.

389 Vojnits K, Pan H, Mu X, Li Y. Characterization of an Injury Induced Population of Muscle-Derived Stem Cell-Like Cells. Sci Rep 2015; 5: 17355.

390 Voorham QJ, Janssen J, Tijssen M, Snellenberg S, Mongera S, van Grieken NC *et al*. Promoter methylation of Wnt-antagonists in polypoid and nonpolypoid colorectal adenomas. BMC Cancer 2013; 13: 603.

391 Wales MM, Biel MA, el Deiry W, Nelkin BD, Issa JP, Cavenee WK *et al*. p53 activates expression of HIC-1, a new candidate tumour suppressor gene on 17p13.3. Nat Med 1995; 1: 570-577.

392 Wallmen B, Schrempp M, Hecht A. Intrinsic properties of Tcf1 and Tcf4 splice variants determine cell-type-specific Wnt/beta-catenin target gene expression. Nucleic Acids Res 2012; 40: 9455-9469.

393 Wang J, Kumar RM, Biggs VJ, Lee H, Chen Y, Kagey MH *et al*. The Msx1 Homeoprotein Recruits Polycomb to the Nuclear Periphery during Development. Dev Cell 2011; 21: 575-588.

394 Wang J, Abate-Shen C. Transcriptional repression by the Msx1 homeoprotein is associated with global redistribution of the H3K27me3 repressive mark to the nuclear periphery. Nucleus 2012; 3: 155-161.

395 Wang J, Abate-Shen C. The MSX1 homeoprotein recruits G9a methyltransferase to repressed target genes in myoblast cells. PLoS One 2012; 7: e37647.

396 Wang JY, Wang CL, Wang XM, Liu FJ. Comprehensive analysis of microRNA/mRNA signature in colon adenocarcinoma. Eur Rev Med Pharmacol Sci 2017; 21: 2114-2129.

397 Wang S, Krinks M, Lin K, Luyten FP, Moos M, Jr. Frzb, a secreted protein expressed in the Spemann organizer, binds and inhibits Wnt-8. Cell 1997; 88: 757-766.

398 Wang S, Chen YG. BMP signaling in homeostasis, transformation and inflammatory response of intestinal epithelium. Sci China Life Sci 2018; 61: 800-807.

399 Wang W, Chen X, Xu H, Lufkin T. Msx3: a novel murine homologue of the Drosophila msh homeobox gene restricted to the dorsal embryonic central nervous system. Mech Dev 1996; 58: 203-215.

400 Watanabe K, Biesinger J, Salmans ML, Roberts BS, Arthur WT, Cleary M *et al*. Integrative ChIP-seq/microarray analysis identifies a CTNNB1 target signature enriched in intestinal stem cells and colon cancer. PLoS One 2014; 9: e92317.

401 Wehkamp J, Stange EF. Paneth's disease. J Crohns Colitis 2010; 4: 523-531.

402 Weinberg RA. Oncogenes, antioncogenes, and the molecular bases of multistep carcinogenesis. Cancer Res 1989; 49: 3713-3721.

403 Whyte JL, Smith AA, Helms JA. Wnt signaling and injury repair. Cold Spring Harb Perspect Biol 2012; 4: a008078.

404 Wielenga VJ, Smits R, Korinek V, Smit L, Kielman M, Fodde R *et al*. Expression of CD44 in Apc and Tcf mutant mice implies regulation by the WNT pathway. The American journal of pathology 1999; 154: 515-523.

405 Willert K, Brown JD, Danenberg E, Duncan AW, Weissman IL, Reya T *et al*. Wnt proteins are lipid-modified and can act as stem cell growth factors. Nature 2003; 423: 448-452.

406 Wilson CL, Heppner KJ, Labosky PA, Hogan BL, Matrisian LM. Intestinal tumorigenesis is suppressed in mice lacking the metalloproteinase matrilysin. Proc Natl Acad Sci U S A 1997; 94: 1402-1407.

407 Wodarz A, Nusse R. Mechanisms of Wnt signaling in development. Annu Rev Cell Dev Biol 1998; 14: 59-88.

408 Wong VW, Stange DE, Page ME, Buczacki S, Wabik A, Itami S *et al*. Lrig1 controls intestinal stem-cell homeostasis by negative regulation of ErbB signalling. Nat Cell Biol 2012.

409 Wong WM, Mandir N, Goodlad RA, Wong BC, Garcia SB, Lam SK *et al*. Histogenesis of human colorectal adenomas and hyperplastic polyps: the role of cell proliferation and crypt fission. Gut 2002; 50: 212-217.

410 Wood LD, Parsons DW, Jones S, Lin J, Sjoblom T, Leary RJ *et al*. The genomic landscapes of human breast and colorectal cancers. Science 2007; 318: 1108-1113.

411 Xu Q, Wang Y, Dabdoub A, Smallwood PM, Williams J, Woods C *et al*. Vascular development in the retina and inner ear: control by Norrin and Frizzled-4, a high-affinity ligand-receptor pair. Cell 2004; 116: 883-895.

412 Yamamoto H, Itoh F, Hinoda Y, Imai K. Suppression of matrilysin inhibits colon cancer cell invasion in vitro. Int J Cancer 1995; 61: 218-222.

413 Yamamoto H, Awada C, Hanaki H, Sakane H, Tsujimoto I, Takahashi Y *et al*. The apical and basolateral secretion of Wnt11 and Wnt3a in polarized epithelial cells is regulated by different mechanisms. J Cell Sci 2013; 126: 2931-2943.

414 Yan D, Wiesmann M, Rohan M, Chan V, Jefferson AB, Guo L *et al*. Elevated expression of axin2 and hnkd mRNA provides evidence that Wnt/beta -catenin signaling is activated in human colon tumors. Proc Natl Acad Sci U S A 2001; 98: 14973-14978.

415 Yan KS, Chia LA, Li X, Ootani A, Su J, Lee JY *et al*. The intestinal stem cell markers Bmi1 and Lgr5 identify two functionally distinct populations. Proc Natl Acad Sci U S A 2012; 109: 466-471.

416 Yan KS, Gevaert O, Zheng GXY, Anchang B, Probert CS, Larkin KA *et al*. Intestinal Enteroendocrine Lineage Cells Possess Homeostatic and Injury-Inducible Stem Cell Activity. Cell Stem Cell 2017; 21: 78-90 e76.

417 Yang PT, Lorenowicz MJ, Silhankova M, Coudreuse DY, Betist MC, Korswagen HC. Wnt signaling requires retromer-dependent recycling of MIG-14/Wntless in Wnt-producing cells. Dev Cell 2008; 14: 140-147.

418 Yang Q, Bermingham NA, Finegold MJ, Zoghbi HY. Requirement of Math1 for secretory cell lineage commitment in the mouse intestine. Science 2001; 294: 2155-2158.

419 Yingling J, Toyo-Oka K, Wynshaw-Boris A. Miller-Dieker syndrome: analysis of a human contiguous gene syndrome in the mouse. Am J Hum Genet 2003; 73: 475-488.

420 Yoshimoto M, Itoh F, Yamamoto H, Hinoda Y, Imai K, Yachi A. Expression of MMP-7(PUMP-1) mRNA in human colorectal cancers. Int J Cancer 1993; 54: 614-618.

421 Yu S, Tong K, Zhao Y, Balasubramanian I, Yap GS, Ferraris RP *et al*. Paneth Cell Multipotency Induced by Notch Activation following Injury. Cell Stem Cell 2018; 23: 46-59 e45.

422 Yue Y, Yuan Y, Li L, Fan J, Li C, Peng W *et al*. Homeobox protein MSX1 inhibits the growth and metastasis of breast cancer cells and is frequently silenced by promoter methylation. Int J Mol Med 2018; 41: 2986-2996.

423 Yue Y, Zhou K, Li J, Jiang S, Li C, Men H. MSX1 induces G0/G1 arrest and apoptosis by suppressing Notch signaling and is frequently methylated in cervical cancer. Onco Targets Ther 2018; 11: 4769-4780.

424 Zeineldin M, Neufeld KL. Understanding phenotypic variation in rodent models with germline Apc mutations. Cancer Res 2013; 73: 2389-2399.

425 Zeng W, Wharton KA, Jr., Mack JA, Wang K, Gadbaw M, Suyama K *et al*. naked cuticle encodes an inducible antagonist of Wnt signalling. Nature 2000; 403: 789-795.

426 Zhai Y, Iura A, Yeasmin S, Wiese AB, Wu R, Feng Y *et al*. MSX2 is an oncogenic downstream target of activated WNT signaling in ovarian endometrioid adenocarcinoma. Oncogene 2011; 30: 4152-4162.

427 Zhang W, Glockner SC, Guo M, Machida EO, Wang DH, Easwaran H *et al*. Epigenetic inactivation of the canonical Wnt antagonist SRY-box containing gene 17 in colorectal cancer. Cancer Res 2008; 68: 2764-2772.

428 Zorn AM, Barish GD, Williams BO, Lavender P, Klymkowsky MW, Varmus HE. Regulation of Wnt signaling by Sox proteins: XSox17 alpha/beta and XSox3 physically interact with beta-catenin. Mol Cell 1999; 4: 487-498.

429 Zou WY, Blutt SE, Zeng XL, Chen MS, Lo YH, Castillo-Azofeifa D *et al*. Epithelial WNT Ligands Are Essential Drivers of Intestinal Stem Cell Activation. Cell Rep 2018; 22: 1003-1015.

8 Appendices

List of publications

1. Msx1 loss suppresses formation of the ectopic crypts developed in the Apc-deficient small intestinal epithelium. **Horazna M**, Janeckova L, Svec J, Babosova O, Hrckulak D, Vojtechova M, Galuskova K, Sloncova E, Kolar M, Strnad H, Korinek V. *Sci Rep*. 2019 Feb 7; 9(1):1629.

2. Wnt Effector TCF4 Is Dispensable for Wnt Signaling in Human Cancer Cells. Hrckulak D, Janeckova L, Lanikova L, Kriz V, **Horazna M**, Babosova O, Vojtechova M, Galuskova K, Sloncova E, Korinek V. *Genes*. 2018 Sep 1; 9(9).

3. Wnt Signaling Inhibition Deprives Small Intestinal Stem Cells of Clonogenic Capacity. Janeckova L, Fafilek B, Krausova M, **Horazna M**, Vojtechova M, Alberich-Jorda M, Sloncova E, Galuskova K, Sedlacek R, Anderova M, Korinek V. *Genesis*. 2016 Mar; 54(3):101-14.

4. HIC1 Tumor Suppressor Loss Potentiates TLR2/NF-κB Signaling and Promotes Tissue Damage-Associated Tumorigenesis. Janeckova L, Pospichalova V, Fafilek B, Vojtechova M, Tureckova J, Dobes J, Dubuissez M, Leprince D, Baloghova N, **Horazna M**, Hlavata A, Stancikova J, Sloncova E, Galuskova K, Strnad H, Korinek V. *Mol Cancer Res*. 2015 Jul;13(7):1139-48.

5. NKD1 marks intestinal and liver tumors linked to aberrant Wnt signaling. Stancikova J, Krausova M, Kolar M, Fafilek B, Svec J, Sedlacek R, Neroldova M, Dobes J, **Horazna M**, Janeckova L, Vojtechova M, Oliverius M, Jirsa M, Korinek V. *Cell Signal*. 2015 Feb; 27(2):245-56.

6. Monensin inhibits canonical Wnt signaling in human colorectal cancer cells and suppresses tumor growth in multiple intestinal neoplasia mice. Tumova L, Pombinho AR, Vojtechova M, Stancikova J, Gradl D, Krausova M, Sloncova E, **Horazna M**, Kriz V, Machonova O, Jindrich J, Zdrahal Z, Bartunek P, Korinek V. *Mol Cancer Ther*. 2014 Apr; 13(4):812-22.

Publications number 1 and 4 constitute the groundwork for this thesis and are enclosed.